

1 TGCCCCGCTGC CCGCCCCGAG TTCCCCGGCCC CGCTGGCCCC AGTCATGGCG
51 AAGCAGTACG ATGTGCTGTT CCGGCTGCTG CTGATCGGGG ACTCCGGGT
101 GGGCAAGACC TGCCTGCTGT GCGCTTCAC CGACAACGAG TTCCACTCCT
151 CGCACATCTC CACCATCGGT GTTGACTTTA AGATGAAGAC CATAGAGGTA
201 GACGGCATCA AAGTGCAGGAT ACAGATCTGG GACACTGCAG GGCAGGAGAG
251 ATACCAAGACC ATCACAAAGC AGTACTATCG GCGGGCCAG CGGATATTTT
301 TGGTCTATGA CATTAGCAGC GAGCGCTCTT ACCAGCACAT CATGAAGTGG
351 GTCAGTGACG TGGATGAGTA CGCACCAAGA GGCCTCCAGA AGATCCTTAT
401 TGGGAATAAG GCTGATGAGG AGCAGAAACG GCAGGGGGAG AGAGAGCAAG
451 GGCAGCAGCT GGGCAAGGAG TATGGCATGG ACTTCTATGA ACAACTGCC
501 TGCACCAACC TCAACATTAA AGAGTCATTC ACGCGTCTGA CAGAGCTGGT
551 GCTGCAGGCC CATAGGAAGG AGCTGGAAGG CCTCCGGATG CGTGCAGCA
601 ATGAGTTGGC ACTGGCAGAG CTGGAGGAGG AGGAGGGCAA ACCCGAGGGC
651 CCAGCGAACT CTTCGAAAAC CTGCTGGTGC TGAGTCCTGT GTGGGGCACC
701 CCACACGACA CCCCTCTTCC CTCAGGAGGC CCGTGGCAG ACAGGGGAGC
751 CGGGGCTTTG CCTGCTGCT GTCCCTCTCGT GTGATGACCC TATTGAGTAT
801 CAGTAGCCAC TACTCCCCCT GCCTGGCCCT GAGAGCGGCT CTGCTGTCAT
851 CTCAAGCAGC CCCGTCCCCC AGCCCGTCCA CCCTGGAGTG GTCTTCTTC
901 GCCTGTTTCC CCAGGCCACAG GCCTGCTACG ACCCCCACGA TGTGCCGAA
951 GCACTGTC ACCATCCCGC ACCCACCAAGA CAACAGCCAG GGCTGGAGTC
1001 CAGGCCACTT TCAGCTGCTC CTTCCTCGT GCATCGTGT TCTTCCTCTGC
1051 TTTTCTCTC TTCCCCCACT TCTCTTCTC TGACCCCTCC CCTCCGGTGC
1101 GTTCTGTATC AAAGCTCTC AAACCCCGTC CCCCCGTGTGT CCTGCTGTGT
1151 GCAGCTCGCT CTTCCTTCC TTCCCTAAGCT ATCCAAGGGG ATGGACCCAG
1201 GCTCGTGGGG AGGTTCCACC CTTGGATCCA GGAAGAACCC TCCACCCCTGC
1251 CTCGTGGGTG GGCCTAAAGGC TACAGGGTGC TTCTTCCTCT TCCCCCACCC
1301 CCACAGTCCCC TCATGTGCCA TGGGCCTGCC TCCCCAGTGA CCTGCAAAG
1351 TGGAGCATCG AGGTAGGAGG GAAACAGCAA CGGGGAGTC CTCGAGCCTG
1401 GGGCTGCCCT ACCTCTACCC ATTCCCCGAC CAGAGCTTTG CCCITGCTTG
1451 GCTGCCGCC TGGCTCTTTG GGGAACTGAG CTCAGAGGCA GGTGCTTCAG
1501 AGAAGGAAAC AAAATGAGGG GTGGCAGGG AAAAAAGTCA CCTCCATTCT
1551 CTACCTCCCA TGCAAGCATGA ACACAATTC TCTCCACCTG GCTCCCAAAT
1601 TTAAAGATGT GGACCAAGGC CTGTGGGTAC TCCAGGGCA AGGAGAGCCC
1651 TGGGGTCAGT GACACTGTCA GGCAACCAT GCACTCCACA AAGGGGAGCA
1701 TTTGGAAATG AAGGACTAGC TCCTATGTAT CAGGTTAAGA GCAAGGGAGA
1751 GCTGGCCAGG GACAGCAGTT TGCACAGCAG AGGGGAATGT AGCAACAGCA
1801 GGGCCTCCTA GGGCCCATCT TCCATTCTT AGGTAAGAAG AGCAATTCT
1851 CAGACTCCCA GGGGGAGGAC TGAGCCTAGC CTTCAGCAAC CAAGGTTCTC
1901 CTGGGACCCA AAGTTTATGG GAGAAGGGCA AAGACTTCAT GGGAAAGAGAG
1951 AAGGAAGGCC CTGGGTAGAA ACGCTTGTTG CTGTTCTCTT TGGCTTTAA
2001 GACAAAGCGC TCATCTTGCCT CTCTACCTCC TGATAGGCTT GAGGGTTTGC
2051 CAACCCACACT GTGGCTACAG GTGGAGGGAA GAGGACTCCT TCCCTCCAGAG
2101 TGCTATGTTC AGGAAGTTTC TTTAACCCCA TATGGCCCAA GAGTAGCTCG
2151 TAGGAGGCCCTTAAAGACG GAACAAGTAA TTAACTCAGTT CTACTGGGT
2201 TCCTGCCAC CGTCCCAGG TGGCGAGGC CTAGGAAGAG GGTCAATTCTT
2251 AAGCCACACA TTAGCTGCAC TGCCTGGCTG CAGCCAAAAC AAAGAACTGG
2301 GTGTTGAGTA TTCACTCAACT AAGAACAAA ATCCAGGGCA CTCATATGTG
2351 AAGGATAAGA ACCTCACTTC CTTACTCCTC CAAAAAGAAG TGGGGAAAGA
2401 ACCATCAAAC CTTTCTCTCCT GACTTACCAA ACCAGGAAAAA CAGCAGGAGA
2451 GGGTGGCTCA GGACTTAGGG ACAGGGTATA GCTTAGATGG TGGAAAGCAA
2501 AGGAGAGCAG GAAGTTGTAATCACTGGCT AATGAGAAAAA GGAGACAGCT
2551 AACTCTAGGA TGAAGCTGTG ACTAGGCTGG AGTTGCTTCC TTGAAGATGG
2601 GACTCCTGG GTATCAAGAC CTATGCCACA TCACACTGGG GCTAGGGAG
2651 TAGGTGATGC CAGCCCTCAA GTCTGTCTTC AGCCAGGGAC TTGAGAAGTT
2701 ATATTGGCA GTGGCTCCAA TCTGTGGACC AGTATTCAG CTTTCCCTGA
2751 AGATCAGGCA GGGTGCCTT CATTGTCTT CTCTCCTAGC CCCCTCAGGA
2801 AAGAAGGACT ATATTGTAC TGTACCTAG GGGTCTGGA AGGGAAAACA
2851 TGGAAATCAGG ATTCTATAGA CTGATAGGCC CTATCCACAA GGGCATGAC
2901 TGGGAAAAGG TATGGGAGCA GAAGGAGAAAT TGGGATTTTA GGGTGCAGCT
2951 ACGCTCACCC TAAACTTTG GTGGCCTGG GCATGTCTTG AGGCCAGAC
3001 TGTTAAGCAG GCTCTGCTGG CCTGTTACT CGTCACCAAC TCTGCACCTG
3051 CTGTCTTGAG ACTCCATCCA GCCCCAGGCA CGCCACCTGC TCCGTAGCCT
3101 CCACTATCTC CCTGTGACGG GTGAACCTCG TGTACTGTGT CTCGGGTCCA

FIGURE 1, page 1 of 2

3151 TATATGAATT GTGAGCAGGG TTCATCTATT TAAACACAG ATGTTTACAA
3201 AATAAAGATT ATTCAAACC ACCAAAAAAA AAAAAAAA AAAAAAAA
3251 AAAAAAA (SEQ ID NO:1)

FEATURES:

5'UTR: 1-44
Start Codon: 45
Stop Codon: 681
3'UTR: 684

Homologous proteins:

Top 10 BLAST Hits

		Score	E
CRA	335001101587561 /WO200058473 /org=Homo sapiens /taxon=9...	428	e-119
CRA	18000004937398 /altid=gi 464561 /def=sp P35289 RB15_RAT RAS...	423	e-117
CRA	18000005187045 /altid=gi 7498104 /def=pir T33855 hypotheti...	220	6e-56
CRA	18000004929618 /altid=gi 131798 /def=sp P24407 RAB8_HUMAN R...	216	1e-54
CRA	335001098683352 /altid=gi 11422744 /def=ref XP_001482.1 TR...	214	4e-54
CRA	18000005096141 /altid=gi 2317272 /def=dbj BAA21744.1 (AB00...	214	4e-54
CRA	18000004952869 /altid=gi 131848 /def=sp P22128 RAB8_DISOM R...	212	3e-53
CRA	335001098688905 /altid=gi 11432830 /def=ref XP_007682.1 RA...	211	4e-53
CRA	18000004945380 /altid=gi 131847 /def=sp P22127 RAO1_DISOM R...	211	5e-53
CRA	18000005163099 /altid=gi 7705849 /def=ref NP_057215.1 ras-...	210	6e-53

BLAST dbEST hits:

	Score	E	
gi	12333507 /dataset=dbest /taxon=96...	626	e-177
gi	12120217 /dataset=dbest /taxon=96...	377	e-102

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

From BLAST dbEST hits:

gi|12333507 brain
gi|12120217 epid_tumor

From tissue screening panels:

Fetal whole brain

1 MAKQYDVLFR LLLIGDSVGK KTCLLCRFTD NEFHSSHIST IGVDFKMKTI
51 EVDGIKVRIQ IWDTAGQERY QTITKQYYRR AQGIFLVYDI SSERSYQHIM
101 KWVSDVDEYA PEGVQKILIG NKADEEQKRQ VGREQGQQLA KEYGMDFYET
151 SACTNLNIKE SFTRLTELVL QAHRKELEGL RMRASNELAL AELEEEEKGKP
201 EGPANSSKTC WC (SEQ ID NO:2)

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

205-208 NSSK

[2] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 2

1 92-94 SER
2 206-208 SSK

[3] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 2

1 29-32 TDNE
2 104-107 SDVD

[4] PDOC00007 PS00007 TYR_PHOSPHO_SITE
Tyrosine kinase phosphorylation site

101-109 KWVSDVDEY

[5] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 2

1 18-23 GVGKTC
2 136-141 GQQLAK

[6] PDOC00017 PS00017 ATP_GTP_A
ATP/GTP-binding site motif A (P-loop)

15-22 GDSGVGKT

[7] PDOC00579 PS00675 SIGMA54_INTERACT_1
Sigma-54 interaction domain ATP-binding region A signature

11-24 LLLIGDSVGKTC

BLAST Alignment to Top Hit:

>CRA|18000004937398 /altid=gi|464561 /def=sp|P35289|RB15_RAT
 RAS-RELATED PROTEIN RAB-15 /dataset=nraa /length=212
 Length = 212

Score = 423 bits (1077), Expect = e-117
 Identities = 207/212 (97%), Positives = 209/212 (97%)
 Frame = +3

Query: 45 MAKQYDVLFRLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIEVDGIKVRIQ 224
 MAKQYDVLFRLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIEVDGIKVRIQ
 Sbjct: 1 MAKQYDVLFRLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIEVDGIKVRIQ 60

Query: 225 IWDTAGQERYQTITKQYYRRAQGIFLVYDISSERSYQHIMKWVSDVDEYAPEGVQKILIG 404
 IWDTAGQERYQTITKQYYRRAQGIFLVYDISSERSYQHIMKWVSDVDEYAPEGVQKILIG
 Sbjct: 61 IWDTAGQERYQTITKQYYRRAQGIFLVYDISSERSYQHIMKWVSDVDEYAPEGVQKILIG 120

Query: 405 NKADEEQKRQVGREQGQQLAKEYGMDFYETSACTNLNIKESFTRLTELVLQAHRKEL+GL 584
 NKADEEQKRQVGREQGQQLAKEYGMDFYETSACTNLNIKESFTRLTELVLQAHRKEL+GL
 Sbjct: 121 NKADEEQKRQVGREQGQQLAKEYGMDFYETSACTNLNIKESFTRLTELVLQAHRKELDGL 180

Query: 585 RMRASNELALAELEEEEGKPEGPANSKTCWC 680
 R ASNELALAELEE+EGK EGPANSKTCWC
 Sbjct: 181 RTCASNELALAELEEDEGKTEGPANSKTCWC 212 (SEQ ID NO:4)

>CRA|335001101587561 /dataset=GENESEQ /org=Homo sapiens /taxon=9606
 /mol_type=protein /date=08-FEB-01 /length=218
 /altid=derwent_id|B41604 /altid=derwent_ac|B41604
 /def=Human ORF X ORF1368 polypeptide sequence SEQ ID
 NO:2736 /patent=WO200058473-A2 /pat_section=Claim
 Length = 218

Score = 428 bits (1088), Expect = e-119
 Identities = 212/218 (97%), Positives = 212/218 (97%), Gaps = 6/218 (2%)
 Frame = +3

Query: 45 MAKQYDVLFRLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIEVDGIKVRIQ 224
 MAKQYDVLFRLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIEVDGIKVRIQ
 Sbjct: 1 MAKQYDVLFRLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIEVDGIKVRIQ 60

Query: 225 IWDTAGQERYQTITKQYYRRAQGIFLVYDISSERSYQHIMKWVSDVDEYAPEGVQKILIG 404
 IWDTAGQERYQTITKQYYRRAQGIFLVYDISSERSYQHIMKWVSDVDEYAPEGVQKILIG
 Sbjct: 61 IWDTAGQERYQTITKQYYRRAQGIFLVYDISSERSYQHIMKWVSDVDEYAPEGVQKILIG 120

Query: 405 NKADEEQKRQVGREQGQ-----QLAKEYGMDFYETSACTNLNIKESFTRLTELVLQAHR 566
 NKADEEQKRQVGREQGQ QLAKEYGMDFYETSACTNLNIKESFTRLTELVLQAHR
 Sbjct: 121 NKADEEQKRQVGREQGQQKCPQLLAKEYGMDFYETSACTNLNIKESFTRLTELVLQAHR 180

Query: 567 KELEGRLMRASNELALAELEEEEGKPEGPANSKTCWC 680
 KELEGRLMRASNELALAELEEEEGKPEGPANSKTCWC
 Sbjct: 181 KELEGRLMRASNELALAELEEEEGKPEGPANSKTCWC 218 (SEQ ID NO:5)

Hmmer search results (Pfam):

Model	Description	Score	E-value	N
PF00071	Ras family	323.8	8.2e-95	1
CE00060	CE00060 rab_ras_like	211.0	1.8e-59	1
PF00006	ATP synthase alpha/beta family	4.2	2.1	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00006	1/1	10	24 ..	203	217 ..	4.2	2.1
CE00060	1/1	2	165 ..	16	184 ..	211.0	1.8e-59
PF00071	1/1	10	212 .]	1	198 []	323.8	8.2e-95

bioRxiv preprint doi: <https://doi.org/10.1101/2023.09.04.553000>; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a [aCC-BY-ND 4.0 International license](https://creativecommons.org/licenses/by-nd/4.0/).

1 GCTCAAGATT GCACAGCTGG TGAGTGGTGA CACTGGACT GGAACCCAAG
51 TGTGCCTTAC TCCAGAGCCC TTGGCATGCA CCTGAAACCC CATGTAAGCC
101 CACTGTGGAG ACGCGCACCT CGAAATAATG GAATCCACTA CATCAGTTCC
151 TTTAGCTTTC TGTGTAATCA GAGTAGCTAG CAGGCTCGGG ATTTCGCCCC
201 CCGGCTTTTT TTTTTTTTTT TTTTTGAGAC AGAGTTTTGC TCTTGTTGCC
251 CAGGCTGGAG TGCAATGGCG CAATCTCGC TCACCGCAAC CTTCGCCTCT
301 CAGGTTCAAG CAATTCTCCT GCCTCAGCCT CCCGAGTAGC TGGGATTACA
351 GGCACCGGCC ACCACGCCA GCTAATTTC TTATATTTC AGTAGAGATG
401 GGGTTTCAAC ATGTTGGCA GGCTGGTCTT GAACTTTCC CCTCTTATTA
451 TAATTCAAGAC ACTTAACCTG AAATATACCT TTTCAATGA AGTAAATGGG
501 CTTACCACTT TCCTTGACCT ACTATTGAAA AATACATTCT CCATCCAATA
551 TTCAGCTGAA AACAGGTAT GTACATATAT ACTTTTCATT GCTTTTTTTT
601 TTTTTTTTTT GAGACAAGGT CTCCCTCTGT TGCGCAGGCT GGAGTGCAGT
651 GTCATGATCT CGGCTCACTG CAGCCTTCCC CTAATGGGTT CAAGCAATCC
701 TCCCACCTCA GCCTCTCAAG CCTGGGATTA CAGGGAGGCC ACCGTGGCCA
751 GCTAATTTC TTATTTTTT AGTAGAGACT GGGTTCACT ACATTGGCCA
801 GGCTGGTCTC CAGCTCCTGA CCTCAAAGTG ATCTGCCGC CTCAGCCTCC
851 CAAAGTACTG GGATTACAGG CATGAGCCAA CGCGCCTAGC CTTTCATTGC
901 TTTTTAAAGA CCTAATAGGC TAGACTTTGC TCTCCCTCAA TACTCGTTGG
951 TAGGGATAGG CAATTCTC AACTCCGGAG AGCATTCTT TGCCCTCTC
1001 CGGTGCTAAC ACATTCACTG GTAGGAAACT GGATCTGAA CAAGGGCCAT
1051 TCATTCTTGC GTGCCACTGG CTATACCACA GAGAAATTAA GGGGTCTGAA
1101 ACAATACATT GGTACACCTGG GCACCTATCC TAAGCACCTT AGAGGGAAAA
1151 CGGAGACTTG CCCGCACACC TCTAAAGGAT TTTGCACTTG GAGATGTTCT
1201 TATGGCCATC TATCTTTCA CCCTGGTGG AGGCGTGAAT AGGCATTTC
1251 CCCATTAAA GAAAAAATGG GGACGGGGGA GGGCGTGAC ACAGTCACAC
1301 AGGTAAGGGG CAGCCAGATG GCAGGGAGGG GGAATTCCAC CCACACTCTC
1351 GGGGACTCAT GGAGACGGGT GTTCGAATCC AGATCTGCT CAAGGCCTTC
1401 CTACCTCGGG TGAGCCCAGC TGAGGTACCA GCCACTGGGG AGCCCGGCCA
1451 GATCCTGCAG ATGCAGGGTG CCACGGGGGG CGGAATTACC GGCAGCCAGAC
1501 TTGGGGTGGG ATATGGGGAG AAGTGGTGAG CCCGGAAAGC GGAGCACGGT
1551 AGAAGTGGGC TGGGTGGGGG CTCACCTCAA CTCCCCCATT CGGAGCGTCC
1601 GCGGAAAAAC GAAAACGTTG CCCCAGGGGG GGCAGGAAGG GGTTGGGAGG
1651 GGGGGCTGGC GCCCCGGCCC AGCGTCGCT GCTCGATGGG GTCCCGCTCT
1701 CCTGCGCGCG CTCCCCGGCC CCTCTCTACC GGGCGGCGG CGGCGCGC
1751 GGGGAAGGGG CGGGCAGGGG CGGCCGCCGG TTTCTCCTCC CACCGCCTCG
1801 CGCCAGCCCA GCGAGCCGA GCGAGCCGA GCGGCGCCG CGCCGGGCTC
1851 CCGCCGCAGC CGCGCTTCCC GGCACCCAGC GAGCGAGTGG GCAGGCGGGC
1901 GGGCGAGGCA GCGCGGGGG CGGGGCCCG CGTCTCCTC GCGCCCGCA
1951 GCGTCCCCGG GCGGGCGCG GCGCGATGG CAGCGCGGA GCAGGGCTGA
2001 GCCCAGCTGCC CGCCCGCAGT TCCCGGCCCT GCTGCCCCA GTCATGGCGA
2051 AGCAGTACGA TGTGCTGTT CGGCTGCTGC TGATCGGGGA CTCCGGGGTG
2101 GGCAAGACCT GCCTGCTGTG CCGCTTCACC GACAACGAGT TCCACTCCTC
2151 GCACATCTCC ACCATCGGT AGGGGCGGTG GCCCCGGGGC CCCCTCCCTC
2201 CCCGCCCGCG GCCCCCTTCC CGGCCGCCCG CGTCCCCAGC TGGGGAGGAA
2251 TTGCCAGCCC CTCCGGCTGG AGGCGGTGGC GCGGAGGGCC GGAGTCCGGG
2301 ATAAATCTCG GGGTGAGCAT AGGTTTGGC AGGTGAGGGT GTCCCTGCTG
2351 CCTGCCGTCC GGACCGAGGG TGGGGTCTCC CGCCTCTTGC CGGGAAAGCCT
2401 TCCGTCCTCAT CAAACCGAGA AACCGGGGT GAGGGGAGCT GGTGTAGGCC
2451 TGGGTACCCC GAGCTGGGGT AGCAAGAACAT GTAGCCGCTG GAATAACACC
2501 CCCACACCCC CAGGGGGAGG GGAAGTAAAG CTTCTGCTAC GGAAAAGGGG
2551 GTCAGGGTGG AGACCGAGT CACTGAGGCC CCCTGGTTC TGTGGTGACC
2601 CAAGGTGGAG CGGGCGGGGG GCGAGGGGG GAAGAGAGGA CGTACGGAGG
2651 GGCCACAGGG ATCGAGTTTC CAGGGCAGAG TTGGGAAGGT AAGCCGCAAG
2701 GTGGGACACC TGGGGGAGGA CACAGATAGG GTGAGGGAGCC CCTGCGCCTG
2751 GGAAGAGGAG ACATCTGTT TGAGGGAGGC TAAAGAGGAT GGAGGAGTGT
2801 CAGGAATACC TGCCCCAGACC AAGGGGTCA GAGGCAGGCCA GGACCCGCC
2851 GAGGGCATCT CTCATCTGGC AGTGCTGGAG CCTGTGGTTA GAGGGACAAG
2901 ACCCGGTGGC ATCCCAGACA GCACTATGAT GGGGTCACTT ATTCTAGGAA
2951 TGGGTCCATG GCCTCCCCCTC TGAGACAGTC AGTCTCCCGC TTCTAGGCTG
3001 TGAGGGGCC CCTCCCTGAG AAGTCTGAGT AGAGGAAATT TCATCCTCAG
3051 CTGCTACCCG GGTCAAGCCCT GGAGTAGGCT CTGCAATTGCC CAAGCCCCCTG
3101 GAAACACCTG CTGGCTGGCT GGTCACTCCAT TTGGAATGCT CTCCTAGAAG

FIGURE 3, page 1 of 14

3151 TCCCTGCTGC CATCAGGGAT GGGCACCAGC TCTCAGCTTC CTCTTGAGGA
3201 TTCATGTCCA CACCATCCCC CCTCCCCCA ACACACATTG CTTGCTGAGA
3251 GAGAAGTAGG AGCAGATAGA TACAGCCAGG AGGAACAGAA CCTTCTGGTT
3301 AAGAAGCCAG CTTTATTGTC CAAGAGACCT GAGACCTCAC TGTGGGGCAA
3351 AGCAACCTTG AATATTGCCT AAACCTCTGA GCTTATTITA GTTCTCATC
3401 TGTAGAACGG GTATAATAAT TGACACCTACC TGCCAAGTTG TTGTCAAGAT
3451 TAAATGAGAT AACGATTGTT AAGTGCTTAG CACAGCCAGA CACATGGTGA
3501 AGCTCGATAA ATGCTGATTG TTCTTACTGC TATTGCCATT ATCATTGAGC
3551 TTTAGGGTC TCCTCTCTT GTTCACCAA CTTGAAGGGT GAAACAACAG
3601 GACTTAGGGT CAGGGAACAG AACTTGTCCG TCTTCTCAG AGGAGCTGTA
3651 AGGCCAACTC TTAGGAAACC CAGGAGCTTG GGCTGAGCCA TGGTTTGGAT
3701 GAGAGACATT GCAGAAAGAA GGGGAGCCTA TAGACACTAA GGCTTTGTGC
3751 CTGCCGGAG GACTGGGGAG AGAGGCAGGT GCAGGAGAAA GGCATGGCG
3801 TGATGGAGGA AGTGGCAGAG GAACCAGATG GTGTATGAGG ACAGGTTG
3851 GGCTCAGGG AAAAGGGCGG TGGGTTATCC CTTAAGGAAA CTAGGAGTGG
3901 CTATTGTTGG GAGAGGCTG GTGCTTGAA CTACTGAGCT ATCTCCAGAG
3951 AGCTGTGGGC TGCCCTGGGAG GCCCTGGCTT TGGCCTGAAG AGCTGTTGTT
4001 TGCACCTGCT CTCCCTAGTCC CATTCCAAGT CCTATAGGTG ACATGGACTT
4051 TTCCCTTTGA GGGCTTCATT CAACCACCTC ATTTCAAGAAG CTCTGGACT
4101 CCTGCTTAGT GCTGTGGGAG GCAGCCTCCC CTGGGAGACA CATAACCTCC
4151 TTTTGAGGG CACCCCTCTT TCTAAAATAC CAGGATGGCC CTCTGAGGCT
4201 CGTGCTCTCC TAAAGAGAG TCCATTGGCT CACACCTCTA ATCATCCACC
4251 CTTCTCTTG TCCTCTTCCCC TTGTAATCTC CCTTCTTCTAGA CACCTTCTGC
4301 TAATAGGTGA ACACAAATAA GGTACAGGG ACTTCCTGAA ACCCTCCAGG
4351 GCAGACCACT TTGGGCACAT AGGTGAATCA GTGAACGTAG TAGGGGTGTC
4401 TCTGCAGCAC TGTCTCCCC CAAGGCCCTT GGTATATTGG CCTAAACCT
4451 AAAGATGGCT CCCAGATTTC TTCCCTCGCT TCTGACACCC GGGTTCCCC
4501 TTCTACAGGA CACAGAGGAT TCTCTAGGGT CCCCCCTTTCC ACAGGACACA
4551 GAGGACTCTA GGAGTTGGA TTCCATGAA TAGAAAGAAA CCTGTCTTC
4601 TTCACACCAG CCTTTTAAAAA TCTGCCAAC TGGGTATCTT AAATGTTTC
4651 TTATTAAAG CTTATTAAGG GACTTGGGAT TCTCCCTTAT CTTGGCGTG
4701 TTTTCAGCA TTAACTAAAA CTTAAAGGAA AGAGTTGGAT GGTCAAGAAA
4751 AGCTTTTCC TTAAGTGATA TGGACAGTT CTCAAGGAGG TAGAAGGGGC
4801 AGCCAGGAGA CAAATCAAGG AGCCAACGAA ATGAGTGCTA CCAAGTCATA
4851 GTCATTGCT TATTTTAAA AAATGCGTGT CCTGTATGCC AGGCTCTGCA
4901 CTGAGACCGA GAGATTCCAA GATGAATAAT ACCTACAGTC ACTGTTCTCA
4951 AATTGTGCAT TACCTAAAAC ACATTACATG ACCATGCTGG CCACTGATCG
5001 AGGCACCTTT CCCAGGGGCT TTTTTGTGA ATTAAGAAAA CAAGGTAAATT
5051 CACCAAGTTAT TGCCAAGATA GTTGGCTTC TTGGCTCATG TGGATATCAC
5101 CTAGGCCAGT ACTTTGTGA TTTACTGTGT ACTCCACTT AACGGCCCTGC
5151 GATCTCTAG AGAAGAACCC GCCAGGGAGC AGTGAGAGGC CTCCCTGGTA
5201 GACTGAGACA CTGACTGTCC CTCCCCCTAT CTTTTGCGTC TTTCTGGCCA
5251 GCAGACCGAG AGGTGGCCCT GCCACTGGCT CTGCCACAGG CATTTCCTT
5301 CTGTCAGCT GTGCTGGCCT GGCTGGGGT TGGTGCAGG GGGTCCCCAA
5351 AATACTACCT TAAACAAATT AATTGAGCAT TCACTACCAA GCTCTGTGCC
5401 AGGCATTTA GAGACATATT GCAGTCTACG TTTCTGCC ACAGAAGCCC
5451 ATAACCTAGA TGGGGAGGCA AGACAAAGGG AAAAACAAAA ACAAAAGAGC
5501 TAGTGCCAAA ATGAGATATC TGAAAGAATC TGGTGAATCA CTCTTCAAAT
5551 GTAAAGGATG GATTATGATC ATTGCAGTTA CTCTTAATGA AGGTCTCACA
5601 GTGGGTATCA GAAGCTAAAT TATGATGCCA GATGTACCAT GAGGCAGCCG
5651 GAGAATGGCG ATGGATGGGA TGGGTGAGTG CTATCCCAC GACTCCATGC
5701 TGTCGGAGGC TGGGGAGAG AGAGGCCCT GTGGACTAGA ACCGGCAGGG
5751 AAGGCTGAAG CTAGGCCTCA GTGTGGCTG CTCGTCAGTT CCTGCAGCAG
5801 AAGGGAGCAG ATGGAGTAAC ATGAGCAGAG ATAACAGAGG TGGGATTGAG
5851 TAGGTGTCCG TGGGGCTCTA GGCAGTTAG ATGCAACAGA AGGGATTCTT
5901 CAGGAAAGTG AGAAGATTCT TCTGTTCTC TCTCTGTCTC CCAAATTATA
5951 AGTGCCTTGA TGGTGCAGC AAATCTTATT CCTCATTGTT TTTATAGTCC
6001 CTAGTACAGG GCCAGGCAGA TTCAATGCC TTTGTTAAAT TAATGAATGA
6051 ATGCAGGGAC CAGTTGGCAG AGGGCATGTA GAGCCTGGCC AAGGAGGTGG
6101 AACATGAGCC TTAGCAATGG TAGGAGGGT TTTGAGTAGG GTACTAATGA
6151 GGTTGGCTGG AAAAAGGGGT TAAGACTGTA AGCAGGGAGA CTAGTCAGGG
6201 GCTGCAGTAG TATCCTGGGC ATGAAGGAAC CTCTGAATGG CCCCTCACCC
6251 CCAGTGGTAC CAACACCAAC TTCCACACAG TCAGTTGTTT TACTTTCCCT

FIGURE 3, page 2 of 14

6301 CCAGAATGGG GAGTGGTTCA AGCCAATCAA CCTGGCAACT TCTGAAAGAA
6351 TCTTATGGGA CCTGTGCCAT GACCAGGTAG GGAGAAGATG TCATACATGG
6401 ACATCTATGT TCAGGGGACC TTTGAGGACC TTTCTGCATG GTGGCCAGGA
6451 GTGAGATGAT GTAAACCACA AATGGAAACT GAAGAGACTG CTCAGGAGTT
6501 GTTGGTTTTC TTTTCTTTTC TATTTTTTTT TTTTGAGAC TAGGTTTCAC
6551 TCTGTACCC AGTCTGGAGT GTGGTGGTGG CACAATCACG GCTCACTGCA
6601 GCCTCGATCT CCTAAACGCA ATCCTCCAC CTCAGCCTCT CAAGTAGCTG
6651 GGACTACAGG TGCATGCCAC CACATTCAAG TAATGTTTGT ACGTTTTGTA
6701 GAGATGGGGT TTCACTATGT TGACCAGCT GGTCTCGAAC TCCTGGACTC
6751 GTGATCCACC AGCCTCAGCC TTCCAAAATG CTGGGATTAT AGGCGTGAGC
6801 TACCTCACTC CCTCAGGAGT TGGTTTCTC CCTCCCATCC TTAGTCTTC
6851 CTGAGTAGAC CTGTCACCTA GTCCCTGGAC CTTTGTGTTT GAAAGCCACC
6901 CTCCAGGCTA CACTCCTTCT GGGTGAGGAG GAGGGTGATC TGGTTGGACA
6951 GGTTGGGCTG CTGTGGCTC AGGGCACTT CTCAGGCTGG GTTGTGCTG
7001 CTATGTCACC TTTCTCAAGG AGTTCTGCTG GGACTGGCTT GGCTGCCTGT
7051 CTTGACTTTG CTTTTGACTG AGGAGGTGGG AGATGGTGAG GGAGGGGGTG
7101 GGGCTAGATC CAAGCCTGGA ATGGGGTGAC CTAACAGACA CTGGGGCCTG
7151 TGCTTAGACA CTAGGATCCT GGGGTTGCA GGTTCTAGA CTGAGAGGAG
7201 CTGGGGCAA ATGCACTGTG ACGTTGTGAG AGGGTCAGGG CTGGGTCTGT
7251 GTCAGCCCTC AGGCAGCCTG AGACCAGTCT CTACCTACTC TGTTCCCCTG
7301 GTACCTAGAA AGGAAGGGAA GGTGAGAAGC AATGAGCAGA ATGAAAGAG
7351 CCCAGATTAA CATGCACATT TCCCAGGCC TTACTGGCCC TGTGACCTTC
7401 AGACACTTG ATGACATCTT TGTGCTTCGT TTCTGCATCT GTAAATTGAA
7451 GATGGTAACA GAGTCTTCT TAAAGGTGTT GTGAAAGATT ATAGAGCCTA
7501 GCGCATATAA AGCACTTGGC AGAGCCCTCG ATAAAATAAT AGCTGCTATC
7551 ATATTATCAT TATTATTATT TTATTTATTT ATTATTTTAT TTTTTTTTGA
7601 GACCGAGTAT CTCTCTGTG CCCAGGCTGG AGTGCAGTGG CACAATCTG
7651 GCTCACTGCA ACCTCCATCT CCCGGTTTA AGTGATTCTC CTGCCCTCAGC
7701 CTCCTAAGTA GCTGGGATTA CAGGCACCCA CCACCCACACC CCGCTATTAT
7751 TATTATTCTC AGCTATAAGA ATGCTGTAGA GATGAATACA CTGTCAGTGA
7801 GCTAGGAGGT CATCCTGTGT ATCCATCACT TGTGCACTCA GTCTGTTAGG
7851 CGCTATTTCG TGAAACACCAA CTACATGCCA GGTGCCATGC TAAGATTTGG
7901 GGACACAGTG GTGACCAAAA CAGACAGAAA CCAAGGAGCT GGCTTACATT
7951 CCAAGGGAGT GCATAGGAAG CTGTGTTCA TTTCAGTTTC TGCTCTAGTA
8001 CCCCCCTTC CCTGGCAGTG CCAGGGCTTG AGAAGGAAGA GTGAGGTGGT
8051 GAGGAGGTGT GAAGCAGTGG GGTGACCTGA GAGGAGAGGA TGGGGTGGCT
8101 TTGCCTCAAG GCTTGGGCCCT CGCTAGGTG TCGCTCTGCC TCAGGCCTCT
8151 GTTTCTCCCTC CTGACACAGG CACAGACTCG GCCTCCACC CCTTCCCCAA
8201 GGACATGACC TTGGGAAGGA ACATATCTGA AGCCCGCGGA GGGTTTCCGC
8251 TGCTGTGCAT CTGTGCCACA GATCCGCAGA TGCAACCCACA GCTGGGAGCA
8301 CCGGTTCCCTC CCGCCTACCT GCACTCCCTG GTTCTGTTC CTTCCCTCC
8351 CTCCCTCCCT CTCCTCCGCTC CCCAGACAGG CTGGTGATGA GCTTTATAAC
8401 ATGAAAGCTG ATATTGGCC ATTATCCTTC TACCCCTGATT GCCAGCTCTT
8451 CTCAGAGTGC CTCTCTCTG AATCCAATCT TTGACCCAGT TTCCCTGTGA
8501 AACTGCCAGT TTCTGTATA GGCCTCTGCC CTCTCCCTGG CTCTTCTCTC
8551 TGGTCAGTGA GCTTTGTCAA GGGGAACACA GGGCTTCCCTG GACACGTAAT
8601 TCCTCCCACT GAGGAGGAAG GGGCTAATCA CCAGCCCTGT TTTATTTTAT
8651 TTTATTTTTTGAGATGAAG TCTAGCTCTG TCGCCCAGGC TGGAGTGC
8701 ATGGCTCGAT CTGGCTCAC TGCAACTCT GTCTCCCGGG TTCAAGCGAT
8751 TCTTCTGCCT CAGCCTCTG AGTAGCTGGG GATTACAAGC ATGCACCACC
8801 ACACCTGGCT AATTCTTGTG GTTTTAGTA GAGATGGGGT TTCACCATGT
8851 TGGCCAGGCT GGTCTCGAAC TTCTGACCTC AGCTGATCCA CCCACCTCGG
8901 CCTCCCAAAG TGCTGGATT ACAGGAGTGA GCCACCATGG CTGGCCGACC
8951 CCATCTCTTA AAAAACAAAA AAGAAAAGAA AAGAAAACAA AACAAAAACA
9001 CTTTTAAAT TAACTGATTAGT TGTTGGCATG TGCTGTAGT CCTAACTACT
9051 CAGGAGGCTG AAGTGAAGG ATTGCTTGAG CCCAAGTAGT TGGAGGCCAC
9101 AGTGAGCTGT GATCACACCA CTGTACTCCA GCCTGGGTGA CAGAGTGAGA
9151 CCCTGTCTCA GGAAAAAAA AAAATTACTG AGAACTCTGT GACCATGGCA
9201 CCATGAACTA TAGAAAGGGC TAACAGTGG CTTTGAAATG TGGGTTATGG
9251 CTGGGTGCGG TGGCTCACCG CTGTAATCCC AGCACTTTGG GAGGCCAAGG
9301 TGGGCAGATC ACAAGGTCAG GAGTTTGAGA CCAGCCCGGC CAACATAGTG
9351 AAACCTCATC TCTACTAAAA ATACAAAAAA TTAGCCGGGT GTTGTGGCAG
9401 GTGCCTGTAA TCCTAGCTAC TCGGGAGGCT GAGGCAGGAG AATTGCTTGA

9451 ACCCAGGAGG TGGAGGTTGC CACAAGCTGA GATCGCACCA CTGCACTCCA
 9501 GCCTGGCGA CAGAGCAAGA CTCCATCTCA AAAACAAAAA TAAAAACAAA
 9551 AAAAAGTGGT TTGTTTCTT TTCTTTCTT TTTCTTTTT TTTTTTTTTT
 9601 TTTGAAACA GAGTCTTGCT CTGTCACCAAG GCTGGATTGC AGTGGAGGAT
 9651 CTCAGCACAC TGCCACCTCT GCCTCCCAGG TTCAAGTGTAT TTCCCTGCCT
 9701 CAGCCTCCAG AGTAGCTGGG ACTACAGGCA CGCACCAACCA CGCTGGGCTA
 9751 AGTTTTGTA TTTTAGTACA GAAGGGGTTT CACCATGTTG GCCAGGATGG
 9801 TCTCCATCTC CCTGACCTCG TGATCCGCC ACCTCGGCCT CCCAAAGTGC
 9851 TGGGATTACG GGCATGAGCC ACCACGCCCG GCCTAAAAGT GGGTTATTTT
 9901 CTAATTGCTC TTCCCTGATT AAAATTTCT CTTTGGCCAT CTTTCTCTA
 9951 GATATGTACT GACTTCATTC ATCCATTAT TCGTCTCACT TGCTCATTCA
 10001 TTTTTGCTTT CATTACTCT ACTTTGTTGA ATAATATTAA GTGATCTACC
 10051 TGCTGCCAGG CAGTGAGGT CTGAAGTGA CAGGATGCTG CTTTGCCTC
 10101 TGGGAGCTTA CAGTGTAGCT GGGAACAGA CATCCAAACA AGCAGAATAT
 10151 TATGAAAAG AAATGTCAGG ATGCTTIGGA ATCACAGAGG AGTGAGAAAT
 10201 CCCTCCCGGG GAGGCTGGTG AAGGCTTGA AGAGGAAGTG ACATTTGAGT
 10251 GGAGTCTTGA AGACTAGGCAGA GGATTCTCA GGGGCCCTGG GTGTGGGGGA
 10301 AGCACACATC CTCTCCCTG TAGGAGGTGC TGTGGAGAAC ACCTCCAGT
 10351 GGGCTGCTAC TCTTCAGCCT TGCTGGGGCC AGCTGGAGTG GCCACACCAT
 10401 GGTACACACCA GCTGAAGTTC AAGAACCCCC TTGCCAGGAG ATTGCTTTGC
 10451 TGGCTCTGGG TGAGGGCAGG TGCACTGGGA AGCCCCCTTC TTTCTAAGAT
 10501 GTTTGCTCCT GAGTTCTAT GTCCTAGTCT TTTCTCCCT GAACCTTTG
 10551 CTACCACTCA GCACAGCCCT GCTGAGAAG GAGGCTGGAG GAGTGAGTGG
 10601 TCAGTAGCCT GGTGGGTCTT GGCTGCCTCT GTGGTCCCCG CTGGCCTAAG
 10651 TAGCAGGCTT AGGGAGGCGA GACCCAGTTC CAGGGCTGC CAATGGGGAG
 10701 CGAGATGGGG TGGCTGGAGC ACACTGCACA TGTCAACAG GCTCTAGGGA
 10751 GGTCTGTCA CAAGGCAGTG GGAAAAGCAA GGGGAAGACC CAGCCTGGTC
 10801 AACATGGTGA AACCCCGTCT CTACTAAAAA TACAAAAATT AGCTGGGTGT
 10851 GGTAGAGCAC GCCTGTAGTC CCAGCTAACT TGGGAGCCTG AGGCAGGAGA
 10901 ATCACTTAA CACAGGAGGT GGAGGTTGCA GTGAGCCGAG ATCGTACAC
 10951 TGTACTCCAG CCTGGGTGAC AGAGTGAGAC CCTGTCCTAA AAAAAAAA
 11001 AAAAAAAA AAAAAGTGGG GAAGGGGAA ACTGATCCTG ATTATCTACT
 11051 CCATATACTT ACTATGTACC TACTACCTAC ACAGGGACGG TGGGCTTTAC
 11101 GCATGCCATT TATTCACTGT ATAGAGATCT CAGCATCACA TAGGAGCAGG
 11151 GAGTTCTGAA GTTGGCCTTG CTGGCATTG AGAAGTTCT TGGTGTATT
 11201 TTCAGGTTCA CGCCTCCAGA CAAGTGTAA TGCTATTGAA TGCTGACTAT
 11251 GTTCCAGGAA CAAACCCAGA TGCTAGAAGA CACGAGTAA ACAGTACAGA
 11301 TGCAGGTGCA CATGTGAGGG TCCACACAAG ACCTGAGAGA AGGGAGGGGT
 11351 CTTGCTGCAG TTCCCCCTTT GTAACAAAG AGAGAGTACT GTTGACCCCTC
 11401 TTCTTAGGAA CTGTGAGTTT GACTGAAATG TGTCTGCCA CAGGATCTTT
 11451 GCTGCTCCT CTACCTGATT CTTTGGATCT CCCTGCTGGC ACCTTCTTGT
 11501 CATTAGGTC TCAGCTCAA TGTTACCTCC TTTAAAATGT CTTCTCTGGC
 11551 CAGCCAGTCT AAGGTTGCTT GTGCTTGGGG TCTCCTCACT CTCTACTTTA
 11601 TCCCGCAGTT GCTCTTATAC ACATATGGCT CTCTGAAATT AGGTATTCT
 11651 TACTTACATC TGTCTTCCCC ACTAGAATTA AGCTCTGATG ACAAGGATCT
 11701 TTCTGTGCTG TTCACTAGCTT ATCTTCTAGT ACCTGGCTTA GTTCCCTGGCA
 11751 CATTGTAAGC ATTCAATAAC AGTTTGAATG AATGAATTAA CAAATGAAGG
 11801 AATGAATGAA TGCACTTCC TAGAGGACTT CTGTTCTTCC CTGAGGGAAAG
 11851 TTATAGGTCG TATTGGTTTC TTGGGACTGT TTTTGTGTTG TTTGTTTGT
 11901 TTTGTTTTT GAGACAGAGT CTCACTGTAT CCCCCAGGCT GGAGTGCAGT
 11951 GGCACAATCT TGCTCACTG CAACTCCGC CTCCCAAGGTT CAAGCGATTC
 12001 TCATGCCCTCA GCCTCCCGAG TAGCTGGGA TTCCAGGAGC CTGCCACCC
 12051 GACCAGCTAA TTTTTGTATT TTTAGTAGAG ACAAGGTTTC ACCATGTTGG
 12101 CCAGGCTGGT CTTGAACCTCC TGACCTCAGG TGACCTGCCT GCCTCTGCCT
 12151 CCCAAAGTGC TGGGATTACA GGCATGAGCC ACCACGCCCG GCCTGTTTT
 12201 TTTTTTTTTT TAAGACAGAG TCTTGCACTG TCTCCCAGAC TGGAGTGCAG
 12251 TGGTGTGATC TCAGCTCATT GCAGCCTCAA CCTCCTGGCC TCAGGTCCAG
 12301 GTGATCCTCT TACCTCAGTC TTCTGAGTAA CTGGGCCCCAC TGGTATATAC
 12351 CACCACACCT GGCTAATTTT TAAATTTTT GCAGAGACAT GGTCTCACTA
 12401 TGTGCCCCCTG ACTGATCTTG AACTCCTGG GTTCAAGTGA TCCTCACACC
 12451 TTGGCTTCCC AAAGTGTGG GTTACAGGT GTGAGCCACC ATGCCTGGGC
 12501 TTGAGACTGT TAAGATGTG AGGCTGGAGG GAGTGGATGG CCTCACTGCT
 12551 TGAGCCCTAG AGATTCCCTA CTCCAGAGTG CCCTGGCTGC AGAGGTGGCC

FIGURE 3, page 4 of 14

12601 CTGGAGGGTC ACTCCAGCAA CCTGGCTGAG CTGATGGCA TCATCTGATA
12651 CCAGCTCTGA CCTGTAAATA TAGGCAACAT GGACCTTAGT CAAAGCACT
12701 GACCCCTCAT CTCTGCATAT ACCAAAGAAG ATGAGATTTG GGTGAGGACA
12751 CAGCCAAACC ATATCAGCTC CGGGGATCCC TGTGTGAATG GGGTCTTTT
12801 TGGTGTGACA GGGCTGCACA GGGTGACTC TTTAGAGGTG ACCTCCGTGC
12851 ACAACCCACA GGAGGTGCAC ATGGCCCACA CATGCTGGTT TCCTGCAGTG
12901 GGAGGGGCTG GGGCACTCCT GGGACCTGTG CTTGTTAACT GGAGCTGGCC
12951 TGGCCCTGGG GATTGGGTGT CTGCCTTGGG TTTCAGGTGT ATTAGGGTGT
13001 TCCTCGTTGT GGAGTCTCAT TACTAATGAA AAGTCAGGT CGCACTGCTG
13051 GTCCTTGGG CTGTGGTGA TCCTGGTGT AACATTGGC ACCCAGAGGC
13101 AGCCCTGTTT CCACTGAAGC ATGCGGAGCT TGGCTGGCAG GCAGGCAAGC
13151 TGGCAGCTGC CCTTAACCCA TGAGGTGCTG GCCCCTAGT AGGCACACCC
13201 TACCTGTGCC AGAATTGAGG TTGTAGCCAG ACTCCAGGAG CCATCTGGC
13251 CCCACAGGGG GCGGCATTTC CTCTTTTGT TGAAACATTC CAGCCAAGTG
13301 CTGGCTTGGG CTTCATCTCT CTGTCCTACT CTCCCTCCTC TCCCCAACAT
13351 AAGCCCTCTT CTACATCCTA GAGCTTTTC CATTCCCCCT CCTGCAGCTC
13401 TGGGCTCGCT AATCTCATGC TTCCCTAAGG GAGCTTGACG GCTGCTTCTG
13451 CTAACATTAA ATAAAGTTCT GCGTGCCAGA CCCTGTGTTA TGGGTTTAC
13501 ACCTTATCTC ACAATCTTAA AAAAAAAATT CTCTGAGAAT CCTCTGTAC
13551 CCCCCCTTA CAGGTGAGGA AACTGAGGCA AAGATAGGCT AACTGGCTTC
13601 CCCAACACCA TGCAGGTAAT TAGTGTAAA GGCAGGGTTG GAACCAAACCT
13651 TGACCTCCCA ATTGTGCTCT TAATGGCAG GACACTCTGT GTCTTGAGGC
13701 ACACCTCTC CATGTTTCTC AGGGAGGAG ACAGTGATGG
13751 GAAGGGTGT TCTTAGTGT GGATGTGCCC TGCTGCTCC TTTCTGTAAG
13801 CGTCACAGCA CCTCCACTGC TGACTGGGG AGGCACCAAG TTTTCCCTG
13851 TTTGCCACC CAAGGGAGC TAGCTTAGGA GTCACTGAG TGCTGGGTGT
13901 CTGCCTGCT GCATCCCTCT ATCCCTCCCC TGCCCCCGGT GCCCAGAGGA
13951 GGGCCCTGCC TGCTTCTCCA GTTCTCAAC AGCAGCGCTG TCCCAGCACC
14001 CTCGGGCTCC AGTTGTGGCC TGGCAGCTGC TGGGGCAGAC ACCATACAGA
14051 CAGAGTCACA GCAGGAAGAG GATGGGGCCC AGGGCTGCTG CCTCAGGCCA
14101 TGGCTGCATG GCACCATCAG TTGATTGAGG AGCTTTCTT GCCAATGTCT
14151 GAGGCATCAG GTGGCAGGAC ACGTCTCCCT GCTCTTAAGC CTCAGGCATG
14201 CAGCCCTCT TATGCTCTCT TGAGGAGGG GGAGATCCCC CTCATGGAAT
14251 TGCTTTTTTT TTGTTTTTTTG AGGGGCTCT ACAGGGCTCT GCTCTGTAC
14301 TCAGGCTGGA GTGCAGCCTC AACCTCCAG ACTCAAGTGA TCCTCCGTGC
14351 TCAGCCTCCC GAGTAGCTGG GACCACAGGT GGACACCAC ACACCTGGGT
14401 TTTTTGTTT TTGTTTTTTTG GTTTCTAGA GATGGGGTCT CACTTTCTTG
14451 CTCAGCTCTGG TCTCGAAACTC CTGGGCCAA GCAGCTCTCC CACCTCGTCT
14501 TCCCCAAAGTG TTGGGATTAC AGGTGTGAGC CACTGTGCTT GGCTTTTTA
14551 TTATTTTAA ATTGTTTTTG GAATTGCTTC TTATGCTTG GCACTATGCT
14601 GGCACATATGT GGCAGAGATT TTAAAAACGA GCAAAACAAA CAAATGCTT
14651 GTCAACACCA AAATGTATTCTGCCCCCTT AGGTTCTTT TGTGTAGTTG
14701 AGGCTAGAAG AAAAAAATAG GGGCAGTAA GGAGCAGGGAA GCGATGGTTT
14751 AGGAGGTCTT CCTTCAGGCC CCCTGTGTA AGCATCTGGC TCACTAGCTT
14801 GGGGGAGCCA TTAGGCAGCA GTGGCCAATC CTGAGGCACT CTCAGGTGTC
14851 ACTAAGAAA GGGGCATGTG CTCTATGGAT ACCCATGGGC TGAACCTGGAA
14901 GTCTGGCTCG GACCCATGGC TGTGCTAGGA TCCACCGTCC CCAGCCCCAA
14951 CTGCAGTCAG CATGTTCATC ATCCCTAGGC CTCTCCGCTT CTTTCTGCAT
15001 GTTTGTCTGC CTCATGCCCT GCTCATTACCA AACTGGTCAG TCCCCACTGC
15051 CCTGCCTGCCA GTGAGCTGGT TTGATTGGCT TAGCTAAGCT CCCTTGCTC
15101 TGCTGGCCAG GTCACCCCTGT GGGTCACCAAG CAAACCTGTT GATGGTCCAG
15151 TCTGAACCTG CTTCTCCACA AAGAACATGT TGCAACCCAGC CCTGCTTCTC
15201 TGAGCAGAGG TTGGGGCTG AGCTGTTCTA GCCAGAAAGG GACACAGGGT
15251 GTGGCAGGCA CCATGATGGG CATATCTAA GTGCCGGAA AAACAATGAG
15301 CTGCTCTCCG TGCTTTGGGC ACCTGGTGG GAGAGGGCCC ATCTGTCTGA
15351 CTTTCTCCTC CTGGGGCTCT CAGCGCTCTCC GAGAACCTCT GCCAGAGCTG
15401 TGTAGAAGTG TTGTCCTGT TTCTTAACAC TTCTGTGCCCT TATTTCTTTC
15451 TGTACCCAAAG AAAGGAAGTA GACTGTTTG TAGGGACACT GTCGGGGTGA
15501 TGAATCTGGA CTTACTGGAA TCATGAACCA TGCAAGGGAG GAAGGGAGAAA
15551 ATAGGCTATG GTGGGTGTCT TAGTTAGGGC TGGCTGCTGT AACAAAATGC
15601 CTTTAGCTGA GTAAATTAAA GCAAGAGAAA TGTATTGCTC AGAGTTGGG
15651 AGGCTGGAA GTCCAAGATC AGGGTGCCAG CAGATTCACT GTCTGGTGA
15701 GGCTGATGCT CTGTGACAAA GGTGGCACCT TCTAGCTCCA TCCTCACATG

15751 GCAGAAAGAGG GAACAAGCTC CCTCAGACCT CTTTCTAAG GGCGTTAGTC
15801 CCATGCATGA GGGCTCTAAC ATCACGACTG AGTCACCTCC CAAAGCCCTC
15851 ACCTCCCACC AGCACTGCAC TGGGGATTAA GTTTCATAT GGGAAATTTG
15901 GAGGAACACA GACCTTCAGA CCACAGCAGC GGGCTTCTCC TCATGTGCC
15951 CCTGCCTCAC TTCTAGATGC CGCATAATGT CAGTGAAACC CCGTCTCTAC
16001 TAAAAATACA AAAAATTAGC TGGGTGTGGT GGCACTGCCC TGTAATCCCA
16051 GCTACTTGGG AGGCTGAGGC AGGAGAACCG CTTGAACCCA GGAGGCAGAG
16101 GTTGCAGTGA CCTGAGATCG TGCCACTGCA CTCCAGCCTG GGCGACAGAG
16151 GAAGACTCCG TCAAGAAAAA AGAGAAAAGG CATCAGGTAT GCCAGGGTGT
16201 GCGGGAAAAG GCATCGGGTA TGCCAGGGCG TGTGGAAAAA GGCATCGGGT
16251 ATGCCAGGGT GTGTGGGAAA AGGCATCGGG TATGCCAGGG CATGTGGGAA
16301 AAGGTGGTAA GATTCCCTCAG CCTCCCAGGG TTGGGAAGCC TCTGGCCGAG
16351 TGAAGCATAAC CCTGGGTGGG TTTTAAGACA CCAGCTTCC AGTCCAGCTC
16401 AGCTGTGGGA TGTGGGAAACA TGAGTCAGTG GGAACATGAG AATTGGCTTC
16451 CCTGTGGCTC ACAATAATAC CTACTCCTGC CTACTTCATG GGACCCGCAT
16501 AAGAGCTGAG GGATTCCATA GCTCAGGGGT ATGCTGTAAA GACAAGCACT
16551 ATGCACCTGG GTGTGGTTCT GAAACTTTCT TGTGCAGAAG AGTGAGTAGG
16601 GCTGGCGAG TCCTGAGAAT GTGCATTCTCACACACCTC TGATGCTGCT
16651 GATGCTCTAG TCCTCTGGCT GGCAAGGGTA CCTGGTTAGT AGGGGCCAGG
16701 ACTCTGTAAT GCCTTCCACT TCAGGGTTCT CTGGGCTGGT TTTCCCTGACT
16751 CCCCAGGAAG CCTTTATTCA GCAGAGGGAA GGTAGGAGTG AGAGGACTAC
16801 GCTGTCACTG CTCACATAC ATCGTTAAT TTATCCCAGC ACAGCCCTTA
16851 GGAGGGAAGC AGTATTCTCC TTCTACACTT AAGAAAATCG GCCTGGTGGC
16901 GAGGCTCATG CCTATAATCC CAGCACTGTG GGAAGCTGAG GCGGGAGGAT
16951 CGCTGGAGCC CAGGAGTTCA AGACTAGTCT AGGCAATACA GGGAGACCTC
17001 ATCTCTACAA AAAAAAAA AATTAGCTGG GCATGGTGGT GCACACTTGC
17051 AGTCCAGCT ACTTACCCAG AGGCTGAGCT GGGAGGATTG CTTGAGTCCT
17101 GGAGGATCGA GGCTGCAGTG AGCTATGATT GCTCCACTAC ACTCCATCCC
17151 TGGCAACAGA GTGAGACTCC ATCCCCAAA AAAAAAAA TTGAAGCTAG
17201 GAGAAGTTGA GACTTGCCTG AAGTTACACA GTAAGTGCCA GAACCAGGAC
17251 TTGGACCAGG TCTTCTGAC TCCAGGCCA TGGATGTTTC TTCCATGACA
17301 TATATAGCTC TTGAAAATAC TTCTATCTAA TATCACCCAC AGTGCTGTTA
17351 AAAATACAGA TTTCTGGGCC TCACCCCTAA ATTATGATTG AGTAGGGTCTA
17401 GGCACGTCAA GGTCAATTGTT TTTGTCTTTG TTTTAAGTCA CCCCAGGTGA
17451 TTCTAAAGCC GAAGCTCTGC AAAGCACACC TTGAGAAACA GAGAACTCTT
17501 GTGCTCTCGC TCTCTTGACA CTTCAAGGTG AAAACTTTG TCCTAATGTC
17551 GTTCTAAAC TTACGCATGT GTGAGAACCA CTGTGAGAGC TTATTGAAAC
17601 TGATTGCGGG ACCCCATACC TAGAGGGCCT GATTCTATAG GTCTGAGGTA
17651 AGGCCAAGA ATTTCATAT TTGCATTCTG TTTCTTTTC CTTCTTTTC
17701 TTTTTTTTTT TTTTGAGAT GAAGTCTCAC CCTGTGGCC AGACTGGAGT
17751 GCAGTGGCAT GATCTCAGCT CACTGCAGCC TCTGGCTCCT GGGTTAAAGC
17801 GATTCTCCCC ACACCCCCAGA CCCGCTCCCTG AGTAGCTGGG ATTACAGGTG
17851 CCCGCCACCA TGACTAGCTA ACGTTTGAT TTTTAGTAGA GACGGGGTT
17901 TCACCATGTT GGGCAGGCTG GTCTCAAACCT CCGTACCTCA GGTGATCCAC
17951 TCACCTCAGC CTCCCAAGGT CTTGGGATTA CTGGTGTGAG CCACCGCGTG
18001 CGGCCAGAAAT TTGCATTCT AACAAAGTCCC AGGTGATGCT GATGCTGTGG
18051 GTCCAGGGAC ACACTTGAG AACAGCTGT TACTCAGGCG ATATGTGGAC
18101 AGTAGCGTCA TCTTCACCTG GGAGCTTCTC GCAGCATCTC AGGCCTTGCC
18151 CTACACCTAC CAGATCAGAA TCTGCATTAA AACTCAATCC CCGCGTGATT
18201 CTCATGCACC TGGAAAGTTG AGAAATATGA CCTTAGAGGA GCGGAATGT
18251 GAAACCACTG GAGGCAGAGA TAGATGGAGA ATATCTCTTC TTCTCACGGA
18301 TACTAAAGAT GCAACAAAAA GGGCTGACTC TCTGGGTGTG CACCCAGGTG
18351 GGGCTGATGA CCGAAAAGAG GCCAGATGTG GACAGAGGAC TCTTCCCTGA
18401 GGGAAAGGAG AGAGAACTTA GGAAAATCTG AAGAAAAGGAG GTGGCTTCAG
18451 AGGAAAGGA TTCATCTGGG CCATAAAACA GTGGAGAAGG TATCCTGCTG
18501 AGAGCACAGG GGTGGGGAGG GGGTGCCTG GAGCTGAAGT CTTCACTGGG
18551 GGGACAGTGA TAGGTGAACA CACATGTGAA TAAACAGTTT GCTAAGCAGC
18601 TGCGAGGGCT GGGCAAGGTG AGAAAGCATC CGTCTGCAGA GGCCTCAATA
18651 AGGCCAGTGT GTTGACTTTG TCCTGCAGTG CTCAGCAGTG GAAAAAAACCA
18701 ACAGCCACGC AGGGAGAGGG AAGGAGCCAC GATGGGCACG GGTTACTGGG
18751 GCCAGGGCTT GACTGGTAGG TGGACACAGC TGAAGGCCCCA GTTGTGTGG
18801 GAACAGAGCG CAGAAGCAAT AGATTCCCT TGAAGATCCT TGGGCTGTAA
18851 ACCTTTTTA AATTAAGAG AGGTTGTGAG GGCAGGGAGG GAGGAAGGAA

18901 AATCCTTCAG AAGACATAGA CTTACTCTGT TTCTTCCATC ATATGTGAAT
18951 GCATATGAAT AGCCAAAAGG TGAATAAAAC ACATGTTCCC AGGTGGCCAG
19001 TGAGACCTAG GTTGCAAGAT GGTGGGGTGT GTGTGAGGCC GGGGAGTGCT
19051 GCGAGCCCCG GAATTCCCTCA GCCTTAGTCC CCCGCCACAT AGCTAAGAAG
19101 TGAGGGAGGA GGTGAGAAGG AGTCACTGCC CAGCCTCACT TCCGGTGGAG
19151 TACCCCTGTCT CCTTGTCACT TCTGTCTCTG GGGACAGTTG CCTGCTTCA
19201 CCTCTCCCTC CATCCCCCTC TCTCTCACAG GGAAAATTC ACCTTAATAT
19251 TGGAAAGTTCC TCTCCTAGCA AAGTCCTTCT CAGGCACCCA CAGGCAAAAA
19301 GGAAACTAAG CAGAGTTAGG GCTTCCAGGC CTAGCCAAC ACTCGACTCT
19351 CCTCTTGCTT CCCTAAGAAC CAGCGCAAGG GGCAGCGTGG GTTCCAGCAT
19401 AGATGGACCT GTGTTGGAAT CTCTGCACGT GCTGTGCTGA CCCTGGCTAG
19451 CCATTGACCT CTCTGAGCCC TTGTTTCCCT TCCACTAGGC TCTCTGAGGG
19501 CAGGGGCCAT GTCTTTTCA CTGCTCTGTC TGCACTGAGC ACTGTGCAGG
19551 GCACATAGGA AGTTCCCATA AATGTTTGTG GGATAAAAGGA AATAAAACCT
19601 TCTCTCTTCC TGTCCTCCCT GTGATGGCTT TGCACAAGGC ACTGTCCCTT
19651 GCCAGGTTTG CTAGGCTAGT GTGAGGATAA ACCAGGTATA TTACAAATTG
19701 GAGAAAATTT CTCGTTCTTC TTGGAAGAAG GTGCTGTATC ATGAAACAAG
19751 AATGCTTGTG TCCCCTTCTA TGCCAGGTAC TGGGGAGAAA CAGGTGCCTG
19801 ATAACCGTTG ATCCAGGCAG AAATAAGCAT ACTCTCTGCCT CCCAAGGCCT
19851 GATGCTTCTC TCCTCCCTC CTTCCTCCCT CCTCTCTCTC ACTCTTTCTC
19901 TGCAACATGGA AGAAGATGGC TGCCAGGCAT TGCCATTG GAAAAGTACA
19951 GCTCAATGGA TATGAATCAG CTTGGGCAGG CGAGAAATGA TTCACGTCTG
20001 ACCAAATCGA TTAGTTCAAG GTTGGCCCTT CTGCATCTTT TTTCCCTTGT
20051 AATTAAATGA TGATGGTCT TGATGGTGGG AAGGAAGAGA CAGAATTAA
20101 TTTGTTGCCTT TTGTTAGAAA GCTGGGGACA GCACAGATAA GGGAAAGATGT
20151 CTCCCATTG GCAAATAACT GATGCGGAGG TGGAGTGGCA GTGGTGATGG
20201 GGATGCTGGT GCTTCAGGC CTTCGGGCC GGGCAGTGCA GCTGGTGGCA
20251 GACGGTTCGG AACTCTACCA TGTTCCCATC TGAAAAGTGT GGCTGATCAT
20301 GCCCACTCTC GACCTTGCTC CAGGGAGTAC ACAAAAGACGT AAGCTTAATT
20351 AACCCACCAG ACAGTAGCTCT TGAATCCCTG GGCATAGTGC CTGGGTATAG
20401 TTAGAGTTGG GGAGAGGCAT GGTCAGCAAA ACAACCTCCC TCATCTCTC
20451 GTTGTCACTC AGAGTCAGC TGGCTGCTGC TGGGGTGCT GACTTCTCTT
20501 GCTGCAGATT TCTCCAATAT GTTCTGCC TGACAGCATT TGCCAAATCC
20551 CTTCGGTTTC TTGTTGTCTCG TGGCAGCTTA GCTCCTCCAG CCCTGGATG
20601 AAGAACGCTG GAAACTCTTT GCTTCCTTTC CCTCCCGCAG TGACATGCCA
20651 TGCCATGCCA CTGCCTCTTC ATCTGGTCT ATGACAGTCA CTCATAAGCA
20701 CCCGCATGTA CCCGGCCCTG CACTAGCTCA TGACAGCTGC AGTCAATTGG
20751 GCCAGGTGCT GTATCTCATC CGGCCTCCCT AGCAACCCCTC TGAGATACTG
20801 GTAATGTCCC TGATGAAGAT ATTTACTGAG GCAGAAATGG ACGCTCAGTG
20851 AAGCAAGGTG CCTGATGTTA TAGCAATGAG CTATGAGTGG CCAGAGGGAG
20901 GAGATAAGCT CAGGCCTGAC ACCAAAGCCC ATGCTCCCTC TAGTCAACCA
20951 CAGTGCCTCC TATGGTGAAT GAGTGAGTCA GCAACCAAGA CGCATGAGGC
21001 CTTCTTTTG GTGAGCCTTG GCTGGGTGCT GAGGCTTCAG GTACAATCAT
21051 GGGTTGGAAG ACCCCCTCTC TCTCTCCACA GTCTGGCACT ATGACCCCTT
21101 CTGGTTATTAA ACAAGGCAAA GAGAGAGAGG GAAGAAAGCA GGCAAATAAT
21151 GTGGGGTGCT ATTCTCTAGAG ATTAGAATT CAGGAAGGAT AAACACAGCG
21201 TTCTCTCCAG AAGTATAAAAT AGGAAGACTT CACACATGAC TAGAACGAGA
21251 CATGTTTAAAG GTCTGTCGAG TAAGGCAGTG ATGAAGTGA TTTCCCCAGA
21301 TTCACTCTCC CTCCCTCTGGG TCCCCCAGGG CCTTTACTTG TGGCAACTTT
21351 CAGCTCAGGG AGGGAGGAAA GCCCCCTTCA AAGCTTCAGA TACTTCCTTA
21401 AGGTCACTT CTGCTTAAAG AAGGCCTTA CATTACTTCA TCCCTTTGCC
21451 AAATTAACCT GAAAGGAAAC CTTCAAGTG TGATTGCCTG GCCCTTCCCT
21501 GTTCATTTCT CGTGGGTACG CTTCTAATC TTCTTCTTT CTTCCCTTCT
21551 TCAGGTGTTG ACTTTAAGAT GAAGACCATA GAGGTAGACG GCATCAAAGT
21601 GCGGATACAG ATCTGGTGAG CTGGGGAGGA GGAGGAGGCA GATGTAGGAG
21651 AAGAGGACTT CTGGCTGCTC CTTAGCTGCC CCTGCCATGT GTAAAATTCC
21701 TAGGCTTCAC CTGGGATAAC TGGCCACCTC TCTGATGGAT GGAAGCGAAG
21751 TCTCAGAACG CCATCTCTTC CTATAAGCCT TAATCTCAA CCTCTAAAGAA
21801 ACTTTAGGGG ATTGACTACA AGCACCAAAG GGCAGGAATT AGAAGGAAC
21851 GGCACACTAA CCATTGTGAA TTTATCTAG GATTAGGCTT TGCCCTTGGG
21901 CTGTGCCACA CTATGTTAAG ATTGGAAGGA AGGAGGCTAC ACCCCCCATC
21951 ATTTAGGGCG AGACCCCTGAG AGAGTTCTC AGGATAGCAT GATGAAGTTT
22001 CCACAGTAGC AGAGGGTGCT GCTGTGGCTC TCTGCTGAG GTCTTGGAAAG

22051 CACTGCCCTT GCCAGGGTTT AGAGCTCCCT CTCATTCCA CAGCAGTATG
22101 GGCACTGCCT TCAGAGGTCC CATAGGGACT AGGGGTGTAG CAGCATCCCC
22151 TGCCAACTCC CATCCAACCA AATCTGGCA CAGTGGCCAG ATTCCAGAGA
22201 GCTGTCCAAG GCCTGTTCTG GCTGTGGCTT CTGGTTCTG CCAGGAGGGC
22251 AGTTGGCAGG AGGGCCAAG GCCCTGCAGG CCTGGTCAGC ACCAGCACAG
22301 ATGACCAGGC CTCTGACTGC AGATCCCTGT GGGGATCCAA GCATCCCTGG
22351 TTTTCACCC TTTAGCTCCC CAGTTTTCC TACAAGGGGA CAGCTCTGCT
22401 CTTCCCCCTCC CCGTCTGTT CCATGGTCCC TGCTCTCTG AGGGACTGGC
22451 TTTCTCTGC AGGGACACTG CAGGGCAGGA GAGATACCA ACCATCACAA
22501 AGCAGTACTA TCGCGGGGCC CAGGTAAGCC ACCACATTGG GGGTTTCAA
22551 GTGGGAAGCT GCCACCCACA CTCCCAGCTC TGGGTATTTG AGATGTCCTGT
22601 GCCACGATC CCCTAAATAC AGTTCGCTG CTTGGAGGGAG CGCAGGGCGT
22651 CTTTCAGCTG TTCACTGATC ATTTGTCCGT CCATTGTTCA TGGCCCACTC
22701 ACTGCAGGCA GGCCCTGTC CTCACCCCTG ACTTCCACCC TCCATCCTGG
22751 GTCAAAGATC CAGGTCAAAG CATGTGGTGT CTTCTGCTG TAGAGAGTT
22801 TGTGATGGC CTGGGAGGC GCACTGGTGG GGTCTGAGAG AAGAGATATT
22851 TCTGGATGCT GAGCAGGGAG AATGGGAGAG TGGGACCCAA CCTTTAAGTT
22901 TCCACGGCCC CTTCTGGCCC CATGACTGCA CTCTCTCTG GCATATCACA
22951 TCTCTCTATT TCTCTCTCTC TCAGGGGATA TTTTTGGTCT ATGACATTAG
23001 CAGCGAGCGC TCTTACCAAGC ACATCATGAA GTGGTCAGT GACGTGGATG
23051 AGGTAGGAGA TGCCACCTCA CTGCCGGGGT GTGGAGAGGG TGCCTCACCG
23101 GGGAAGGCCA GGCGAGGGCC AGATGGGAG GCAAATGCTT CCAGGAAGCT
23151 TTGCGCTTCCA CAGCCCTGGA TGAAGACCTC TGGTGAGTA AGACATGGGG
23201 AAGAAACCGA AGCTGCCATG CCCTCACTCT CTTACCCCTG CCAGGCCTCC
23251 ACGGCTGTGT CTTCCCGGA AATGAATTAG TTCCAAGTCT TCCCTGTGAG
23301 CAGCTCTTT CCTGAAATCT TGGGACCAAG TGGAGTTGCA AGATTGGGAT
23351 CTAGTCTGG CTCACCAA TAGCTGTGG GCCTGGGAA GCCATTGAA
23401 TCCCTGGGT CCCCAGTCC TGTAGAATGA GGGCTGGACT TACATCCAAT
23451 GTCTTCTCA GCTCTGATAC CAGTGGTCTA ACCCAAGGAA GCACCAGTCT
23501 TAGCCAGAGT GTCTTCTACC CTAAGCTCTC CCCGTGATAC CTTGAGGTC
23551 AGCCATGGCA CTTGGGGAG CCTGGCACCT GCATCCAGTC GGCCCACCC
23601 GTCCCTAGGG CTCTGGAAATT GGTGGTGGGC TGGAGGCAGT GCAGACTCTG
23651 TAGGGAAAAT TGGGGGGGCA GGCAGCACTC ACTGGCTGTT CTGCCCATCC
23701 TTTGTCCCTA GTACGCCACCA GAAGGCGTCC AGAAGATCCT TATTGGGAAT
23751 AAGGCTGATG AGGAGCAGAA ACGGCAGGT GGAAGAGAGC AAGGGCAGCA
23801 GGTAAGTGG AGGAAAAGGC AAGTCCACCC CAGGTCCCTCT GCTGGGCCTC
23851 CAGGGCCAGT CCTGAGCGTG GGGACCTAGG GGTGTGTTCC CCAGTGGCAG
23901 GTCTCTCCAC ACGTCCCCAG CACCCCAAGG CCCTGGGGGA GTGGCCATCC
23951 TCGGAAGGCT TGGTGTCTGG GTTTCAGGAC AGAAGCCCAG AGATTGGGG
24001 TCCATCCAGA AACAAAGACG TCATAGGCAG CAACTCTCCC AAGTCCAGGT
24051 CCCAAATGC AGGATTGCC TCTGCTTAAG AGATCATCCC CGTGTAGTA
24101 ATGAAGGACT TCAAGTTGTC AACCTCTCTC CTGACAGCAT CCAGGCCTAG
24151 CTGCCATGTT ACGGTCGAGA AATGATCTCC CATCCCACCC AACACTCCCC
24201 CACTCTGTC CTTCTTACCC AGGAAAGAGC CAGGGAGGA AATGAGGAGA
24251 CAAAGAGCCA CAGCTGGAGA AGCCATGGGG GCAGAAAGGG TAGGAGGATG
24301 ACGCTGAGGG AATGTCCAAG CATGCAGGGA GACCACCTC CCAGAGAGCA
24351 GAAAGAAATA TTGGTATT TTTTTCTT TCTTCTTTTT TTTTTTTTT
24401 TTGAGATGG AGTCTCGCTC TGTCAACCCAG GCTAGAGTGC AGTGGGCCA
24451 TCTCGGCTCA CTGCAACCTC TGCCCTCTGA GTTCAAGCAA TTCTTCTGCC
24501 TCAGCCTCCC AAGTAGCTGA GATTACAGGT GCATGCCACC ACGCCTGGCT
24551 AATTTTTTTG TATTTTTAGT AGAGATGGGG TTTTGCATA TTGGCCAGGC
24601 CGGTCTCGAA CTCTAACCT CAGGTGATCC ACCTGCCTCA GTCTCCCAA
24651 GTGCTGGGAT TACAGGCGTG AGCCACTGTG CCCAGCCAAG ATTGGTATT
24701 CTGAGATAAG TTATCCACTC AGTCCGTGGA CCTCAAGAGT TTTCCTCTCC
24751 CTTTTCAGTC AATAGCGTTC CATTAGTACT TAAAATGAAA TTGATTGTTT
24801 GGTATAAAAT ATAAGACATG GTCATTGACC AATTGAAAG TAGAGGAA
24851 GCCTACTAGG ATAGTATTAA TTGAGCACTC TATGTGTGGC ACTGTGCTAA
24901 GGCAAGCGCT TTTAAGTGC CGACCCCCACT GAATCATCCC ACAACCATGG
24951 ATGGGAGACA CACTCAGTCT CTTTAACAG AAGATAAACG TGGGGCTTAC
25001 AGAGAATGTA CAACTTGTCC AAGGTACACAC AGCTAGCCAT CAGTGGCAGT
25051 GCTGCTATTG AGGTCTGGGA CTGTGGACT CCAGAGCCCA TGTTTTTAC
25101 GAGGATGCCA TACTGCCACA ATGGATGGTG TCTTTATCTC CTGATATATG
25151 ATTGTGTGTT GGGAGGCGTG GGGTGGCAGC TGGAAAGAATG GAGAGGCATA

25201 TTTGTGGAGG ATCTTCCCCC ATTCTCTGCT ACCCTCTCTT GGAGCTCCCC
 25251 GTCCCATCTG AGAAATTATC TACTCTGAGA AATCGTCACA ACACAGCATG
 25301 GTTGTGAGTG CAGTGGCAGA AGCCTGTGCC TGGGTGTATG GGCCCCCTCCC
 25351 CTGCCTTACT GACTCTCTT CAGAAATGTC CTTCTCTTGC AGCTGGCGAA
 25401 GGAGTATGGC ATGGACTTCT ATGAAACAAG TGCCCTGCACC AACCTCAACA
 25451 TTAAAGAGGT GAGAGCCCTG GTGACCAGGC GCCCCGCTCTC TCGGGCTGAG
 25501 TCCAGCAGAG GTGGGAGGAG GAGCCATAAG ATGGACCTTA TCCCTCAGGC
 25551 CGCTGCAGGG TTGCCAGGG AGAGGAGGAG AACTGGACT AACCTGTGCC
 25601 CTTTGGTTTC CAGTCATTCA CGCGTCTGAC AGAGCTGGTG CTGCAGGGCC
 25651 ATAGGAAGGA GCTGGAAGGC CTCCGGATGC GTGCCAGCAA TGAGTTGGCA
 25701 CTGGCAGAGC TGGAGGAGGA GGAGGGAAA CCCGAGGGCC CAGCGAACTC
 25751 TTCGAAAACC TGCTGGTGT GAGTCCTGTG TGGGGCACCC CACACGACAC
 25801 CCCTCTTCCC TCAGGAGGCC CGTGGGAGA CAGGGGAGCC GGGGCTTGC
 25851 CCTGCTGCTG TCCTCTCGTG TGATGACCCCT ATTGAGTATC AGTAGCCACT
 25901 ACTCCCCCTG CCTGGCCCTG AGAGCGGCTC TGCTGTCATC TCAAGCAGCC
 25951 CCTGCCCCA GCCCCGTCAC CCTGGAGTGG TCTTCTTCAG CCTGTTTCCC
 26001 CAGCCACAGG CCTGCTACGA CCCCCCACGAT GTGCCGCAAG CACTGTCCTA
 26051 CCATCCCGA CCCACCAAGAC AACAGCCAGG GCTGGAGTCC AGGCCACTTT
 26101 CAGCTGCTCC TTTCCTCGTG CATCGTGTCT CTTCTCTGCT TTTTCTCTCT
 26151 TCCCCCACTT CTCTTTCTCT GACCCCTCCC CTCCGGTGCCT TTTCTGTATCA
 26201 AAGCTCCTCA AACCCCGTCC CCCGTGTGTC CTGCTGTGTC CAGCTCGCTC
 26251 TTTCCTTCCT TCCTAAGCTA TCCAAGGGGA TGGACCCAGG CTCGTGGGG
 26301 GGTTCCACCC TTGGATCCAG GAAGAACCCCT CCACCCCTGCC TCGTGGTGG
 26351 GCCAAAGGCT ACAGGGTGC TCTTCCTCTT CCCCCACCCC CACTGTCCTC
 26401 CATGTGCCAT GGGCCTGCCT CCCCAGTGAC CTGCGAAAGT GGAGCATCGA
 26451 GGTAGGAGGG AAACGGCAAC CAGGGAGTCC TCGAGCCTGG GGCTGCCCTA
 26501 CCTCTACCCA TTCCCCGACC AGAGCTTTGC CTTGCTTGG CTGCCGCCCT
 26551 GCCTCTTGG GGAAGTGAGC TCAGAGGCAG GTGCTTCAGA GAAGGAAACA
 26601 AAATGAGGGG TGGCAGGGAT AAAAAGTCAC CTCCATTCTC TACCTCCCAT
 26651 GCAGCATGAA CACAATTTCCT CTCCACCTGG CTCCCAAATT TAAAGATGTG
 26701 GACCAAGGCC TGTGGGTACT CCAGGGCAA GGAGAGCCCT GGGGTAGTG
 26751 ACACTGTCA GCCAACCATTG CACTCCACAA AGGGGAGCAT TTGAAATGA
 26801 AGGACTAGCT CCTATGTATC AGGTTAAGAG CAAGGGAGAG CTGGCCAGGG
 26851 ACAGCAGTTT GCACAGCAGA GGGGAATGTA GCAACAGCAG GGCCTCCTAG
 26901 GCCCCATCTT CCATTCTTA GGTAAAGAAGA GCATTTCTC AGACTCCAG
 26951 GCGGAGGACT GAGCCTAGCC TTCAGCAACC AAGGTTCTCC TGGGACCCAA
 27001 AGTTTATGGG AGAAGGGAA AGACTTCATG GGAAGAGAGA AGGAAGGCC
 27051 TGGGTAGAAA CGCTTGGTGC TGTTCTCTT GGCCTTTAAG ACAAAAGCGCT
 27101 CATCTTGCCC TCTACCTCCT GATAGGCTTG AGGGTTTGCC AACCACACTG
 27151 TGGCTACAGG TGGAGGAAG AGGACTCTT CCTCCAGAGT GCTATGTTCA
 27201 GGAAGTTTCT TTAACCCCAT ATGGCCCAAG AGTAGCTCGT AGGAGGCCCT
 27251 TTAAAGACGG ACAAAAGTAAT TTACCAGTTC TACTGGGGTT CCTGCCACC
 27301 GTCCAAGGT GGGCGAGGCC TAGGAAGAGG GTCAATTCTTA AGCCACACAT
 27351 TAGCTGCACT GCGTGGCTGC AGCCAAAACA AAGAACTGGG TGTGAGTAT
 27401 TCATCAACTA AGAACAAAAA TCCAGGGCAC TCATATGTGA AGGATAAGAA
 27451 CCTCACTTCC TTACTCCTCC AAAAAGAAGT GGGGAAAGAA CCATCAAACCC
 27501 TTTCCTCTG ACTTACCAAA CCAGGAAAAC AGCAGGAGAG GGTGGCTCAG
 27551 GACTTAGGGA CAGGGTATAG CTTAGATGGT GGAAAGCAAA GGAGAGCAGG
 27601 AAGTTGTAAA TCACTGGCTA ATGAGAAAAG GAGACAGCTA ACTCTAGGA
 27651 GAAGCTGTGA CTAGGCTGGA GTTGCTTCCT TGAAGATGGG ACTCTTGGG
 27701 TATCAAGACC TATGCCACAT CACACTGGGG CTAGGGAAAGT AGGTGATGCC
 27751 AGCCCTCAAG TCTGTCTTCA GCCAGGGACT TGAGAAGTTA TATGGCAG
 27801 TGGCTCCAAT CTGTGGACCA GTATTTCTAGC TTTCCCTGAA GATCAGGCAG
 27851 GGTGCCATTG ATTGTCTTTC TCTCCTAGCC CCCTCAGGAA AGAAGGACTA
 27901 TATTGTACT GTACCCTAGG GGTTCTGGAA GGGAAAACAT GGAATCAGGA
 27951 TTCTATAGAC TGATAGGCC TATCCACAAG GGCCATGACT GGGAAAAGGT
 28001 ATGGGAGCAG AAGGAGAATT GGGATTTAG GGTGCAGCTA CGCTCACCCCT
 28051 AAACCTTTGG TGGCCTGGGG CATGTCTTGA GGCCCCAGACT GTTAACCAGG
 28101 CTCTGCTGGC CTGTTTACTC GTCACCCACCT CTGACCTGCTC TGTCTTGAGA
 28151 CTCCATCCAG CCCCAGGCAC GCCACCTGCT CCTGAGCCTC CACTATCTCC
 28201 CTGTGACGGGG TGAACTTCCTG GTACTGTGTC TCGGGTCCAT ATATGAATTG
 28251 TGAGCAGGGT TCATCTATT TAAACACAGA TGTTTACAAA ATAAAGATTA
 28301 TTTCAAACCA CCGGTGTGGC TGCCTGGATG AGTCCTTGGG GGTAGGTCTC

FIGURE 3, page 9 of 14

28351 ACTCAGACCC TGGCAGTGAT GTGGGAGGGA GAGAGGCAGT GCTGGTAGAA
 28401 GCAGCTCCAG AAGCAAAGGC AACAGCAGTA GAGTGACCAC GGAAGCGGCA
 28451 AACATTGTCT TCCCTTCTCT ACCTTCCCTA GTGCCACCTG CAGGGAGGCC
 28501 CAAAGCAAAG CCCC GTTGCCTGCTG CAGGGAGGCC
 28551 TGAAACACAG TTATCGAGAG GATGCTGAAC ATCTATGAGC AGGTTTTAAA
 28601 GCCAAGATGA GTCTCATCTG TTTGTGTGGG TCAGGAACGG GTCTTCCTGA
 28651 AGGCATGAGG TGGGACTGGA TAATCTTCA GATTGTGTAT TGGATAACCTC
 28701 GGGGGAGCAG AGGCAGACTG GGATCTCAGG ACTGCAGGTA TTTCATACTT
 28751 TGGGATATGG AATTGATGGA (SEQ ID NO:3)

FEATURES:

Start: 2044
 Exon: 2044-2167
 Intron: 2168-21554
 Exon: 21555-21615
 Intron: 21616-22462
 Exon: 22463-22523
 Intron: 22524-22974
 Exon: 22975-23052
 Intron: 23053-23711
 Exon: 23712-23801
 Intron: 23802-25392
 Exon: 25393-25458
 Intron: 25459-25613
 Exon: 25614-25769
 Stop: 25770

CHROMOSOME MAP POSITION:

Chromosome 14

ALLELIC VARIANTS (SNPs):

DNA				
Position	Major	Minor	Domain	
206	-	T	Beyond ORF(5')	
4963	C	T	Intron	
8175	G	A	Intron	
10515	T	C	Intron	
13034	T	C	Intron	
13781	T	C	Intron	
14050	A	C	Intron	
14273	-	T	Intron	
17582	T	C	Intron	
17700	C	T	Intron	
18074	T	C A	Intron	
19328	G	T	Intron	
19570	A	G	Intron	
20892	C	T	Intron	
26465	G	A	Beyond ORF(3')	
26472	A	G	Beyond ORF(3')	
28071	C	T	Beyond ORF(3')	
28096	C	G	Beyond ORF(3')	
28403	A	G	Beyond ORF(3')	
28467	C	G	Beyond ORF(3')	

Context :

DNA
Position
206

GCTCAAGATTGCACAGCTGGTGAGTGGTACACTGGACTGGAACCCAAGTGTGCCTTAC
TCCAGAGCCCTTGGCATGCACCTGAAACCCCATGTAAGCCACTGTGGAGACGCGCACCT
CGAAATAATGGAATCCACTACATCAGTCCCTTAGCTTCTGTGAATCAGAGTAGCTAG
CAGGCTCGGGATTCGCCCGGCGC
[-, T]
TTTTTTTTTTTTTTTTGAGACAGAGTTGCTCTTGTGCCCCAGGCTGGAGTGC
GGCGCAATCTCGGCTCACCGAACCTCGCCTCTCAGGTTCAAGCAATTCTCCTGCCTCA
GCCCTCCGAGTAGCTGGATTACAGGCACCGGCCACCCAGCCTAGCTAATTTTTATAT
TTTAGAGATGGGTTTACCATGTTGGCCAGGCTGGCTTGAACTTTCCCTCTT
ATTATAATTCAAGACACTAACCTGAAATATACTTCAAATGAAGTAATGGGCTTACC

4963

TATTAAGGGACTTGGGATTCTCCCTTATCTTGGCGTGTGTTTCAGCATTAACTAAA
TAAAGGAAAGAGTTGGATGGTCAAGAAAAGCTTTCTTAAGTGATATGGACAGTTCT
CAAGGAGGTAGAACGGGCAGCCAGGAGACAAATCAAGGAGCCAACGAAATGAGTGCTACC
AAGTCATAGTCATTGCTTATTTTAAAAAATGCGTGTCTGTATGCCAGGCTCTGCACT
GAGACCGAGAGATTCAAGATGAATAATACCTACAGTCAGTCTCAAATTGTGCATTA
[C, T]
CTAAAACACATTACATGACCATGCTGCCACTGATCGAGGCACCTTCCAGGGGCTTT
TTTGTGAATTAAAGAAAACAAGGTAATTCAACCGATTATTGCCAAGATAGTTGGCTTCTG
GCTCATGTGGATATCACCTAGGCCAGTACTTTGTGATTTACTGTGACTCCACTTAAC
GGCCTGCGATCTCTAGAGAAGAACCGCCAGGAGCAGTGAGAGGCCTCCCTGGTAGAC
TGAGACACTGACTGTCCCTCCCTATCCTTCTGCTTCTGCCAGCAGACCAGCAGG

8175

ATGCCAGGTGCCATGCTAAGATTTGGGACACAGTGGTACCAAAACAGACAGAAACCAA
GGAGCTGGCTTACATTCCAAGGGAGTGATAGGAAGCTGTGTTCAATTCAAGTTCTGCT
CTAGTACCCCCCTTCCCTGGCAGTGCCAGGGCTGAGAAGGAAGAGTGAGGTGGTAGG
AGGTGTGAAGCAGTGGGGTGAACCTGAGAGGAGAGGATGGGGTGGCTTGCCTCAAGGCTT
GGGCCCTGCTAGGTGTCGCTGCCTCAGGCCCTGTTCTCCTCTGACACAGGCACA
[G, A]
ACTCGGCCCTCCACCCCTCCCAAGGACATGACCTGGGAAGGAACATATCTGAAGCCC
GGGGAGGGTTTCCGCTGTCGACTCTGCACTCCCTGGTTCTGTTCCCTCCCTCC
GAGCACCGGTTCCCTCCCGCTACCTGCACTCCCTGGTTCTGTTCCCTCCCTCC
TCCCTCTCCCGCTCCCCAGACAGGCTGGTGATGAGCTTATAACATGAAGCTGATAATT
TGGCCATTATCCTCTACCTGATTGCCAGCTCTCAGAGTGCCCTTCTGTAATCC

10515

CTGGTGAAGGCTTGAAGAGGAAGTGACATTGAGTGGAGTCTGAAAGACTAGGCAGGAT
TCTCCAGGGCCCTGGGTGTGGGGAAAGCACACATCCTCTTCCCTGTCAGGAGGTGCTGTG
GAGAACACCTCCAGTGGGCTGCTACTCTCAGGCCCTGCTGGGGCAGCTGGAGTGGCA
CACCATGGTCACACCAGCTGAAGTCAAGAACGCCCTTGCCAGGAGATTGCTTGTGG
TCTGGGTGAGGGCAGGTGCATCTGAAGGCCCTCTTCTAAGATGTTGCTCTGAGT
[T, C]
TCTATGTCTTAGTCTTTCTTCCCTGAACCTTTGCTACCAGTCAGCACAGCCCTGCC
AGAAGGAGGCTGGAGGAGTGAGTGGTCAGTAGCCTGGTGGCTGCTCTGTGG
GCCCGCTGGCTAAGTAGCAGGCTTAGGGAGGGAGACCCAGTCCAGGGCTGCCAATG
GGGAGCGAGATGGGGTGGCTGGAGCACACTGCACATGTCACCAAGGCTTAGGGAGGTCT
GTGACAAGGCAGTGGAAAAGCAAGGGGAAGACCCAGCCTGGCAACATGGTAAACCC

13034

AGATTGGGTGAGGACACAGCAAACCATATCAGCTCCGGATCCCTGTGTAATGGG
TCTTTTTGGTGTGTTGAGGGCTGCACAGGGTGCACCTTCTAGAGGTGACCTCTGCCACA
ACCCACAGGAGGTGCACATGGCCACACATGCTGGTTCTGCACTGGAGCTGGCCTGG
CACTCCTGGGACCTGTCGGTAACTGGAGCTGGCCTGGCCCTGGGATTGGGTGCTG
CCTGGGTTCAAGGTGATTAGGTGTTCTCGTGTGGAGTCTCATTACTAATGAAAG
[T, C]
TCAGGGTGCAGTGCTGGTCCTTGGCTGTGGTGATCCTGGTGATAACATTGGCACCC
AGAGGGCAGCCCTGTTCCACTGAAGCATGCGGAGCTGGCTGGCAGGCAGGCAAGCTGG
AGCTGCCCTTAACCCATGAGGTGCTGGCCCGCTAGTAGGCACACCCCTACCTGTGCCAGAA
TTGAGGGTTGAGGCCAGACTCCAGGAGCCATCTGGCCCCACAGGGGGCGGCATTCTCT
TTTGTGAAACATTCCAGGCCAGTGCTGGCTGGCTCATCTCTGTCCCACCTCTC

FIGURE 3, page 11 of 14

13781	CCCTGTGTTATGGGTTTACACCTATCTCACAACTTAAAAAAATTCTCTGAGAAT CCTCTGTCACCCCACTTACAGGTGAGGAAACTGAGGCAAAGATAGGCTAACTGGCTTC CCCAACACCATGCAAGTAATTAGTGTATAAAGGCAGGGTGGAACCAAACCTGACCTCCA ATTGTGCTCTTAATGGCAGGACACTGTGTCTTGAGCCACACTCCTCCATGTTCT AGGGCTTCTAGGGAGGCAGACAGTGTGGAGGGTGTCTTAGTGTGGATGTGCC [T, C] GCCTGCTCTTCTGTAAAGCGTCACAGCACCTCACTGCTGTACTGGGAGGCACCAAGT TTTTCCCTGTTGCCACCCAAGGCAGGCTAGCTTAGGAGTCACGTGAGTGCTGGGTGTC TCGCCTGCTGCATCCCTATCCTGCCCTGCCCGGTGCCAGAGGAGGGCCCTGCC GTCTTCCAGTCTCCAACAGCAGCGCTGTCCCAGCACCTCGGCTCCAGTTGTGCC GGCAGCTGCTGGGCAGACACCATACAGACAGAGTCACAGCAGGAAGAGGATGGGCC 14273
	GGAAGGGGTGTTCTTAGTGTGGATGTGCCCTGCCCTGCTCTTCTGTAAAGCGTCACAGC ACCTCCACTGTGTACTGGGAGGCACCAAGTTTCCCTGTTGCCACCCAAGGCAG CTAGCTTAGGAGTCACGTGAGTGCTGGGTGCTGCCCTGCTGCATCCCTATCCTGCC CTGCCCGGTGCCAGAGGAGGGCCCTGCCCTGCTCTCCAGTCTCCAACAGCAGCGCT GTCCCAGCACCTCGGCTCCAGTTGTGGCTGCCAGCTGCTGGGCAGACACCATACAG [A, C] CAGAGTCACAGCAGGAAGAGGATGGGCCAGGGCTGCTGCCCTCAGGCCATGGCTGCATG GCACCATCAGTTGATTGAGGAGCTTCTTGCCAATGTCAGGAGCATCAGGTGGCAGGAC ACGTCTCCCTGCTCTTAAGCCTCAGGCATGCAGGCCCTCTTATGCTCTCTGGGTGAGGG GGAGATCCCCCTCATGAAATTGCTTTTTTTTTTTTTTTTTTTGAGACAGGGTCT GCTCTGTCACTCAGGCTGGAGTGCAGCCTCAACCTCCCAGACTCAAGTGTACCTCTGCC 17582
	TCTCCAACAGCAGCGCTGTCCCAGCACCTCGGCTCCAGTTGTGGCCTGGCAGCTGCTG GGCAGACACCATACAGACAGAGTCACAGCAGGAAGAGGATGGGCCAGGGCTGCTGCC TCAGGCCATGGCTGCATGGCACCATCAGTTGATTGAGGAGCTTCTGCCAATGCTGA GGCATCAGGTGGCAGGACACGTCTCCCTGCTCTTAAGCCTCAGGCATGCAGGCCCTCTTA TGCTCTCTGGGTGAGGGGAGATCCCCCTCATGAAATTGCTTTTTTTTTTTTTTT [T, T] TTTGAGACAGGGCTCTGCTGTCACTCAGGCTGGAGTGAGCCTCAACCTCCCAGACT CAAGTGATCCTCTGCCCTAGCCTCCAGTAGCTGGGACACAGGTGGACACCATCACA CCTGGGTTTTTTGTTTTGTTTTGTTTTCTAGAGATGGGGCTCTACTTCTGCTC AGTCTGGTCTGAACTCCTGGCGCAAGCAGTCTCCACCTCGCTCTCCAAAGTGT GGATTACAGGTGTGAGCCACTGTGCTGGCTTTTATTTATTTAGAATTGTTGGAA 17700
	GGATGTTCTTCATGACATATATAGCTCTTGAAGACTACTTCTATCTAAATACACCCACA GTGCTGTTAAAATACAGATTCTGGGCCAACCTCAAATTATGATTCACTAGTAGGTCTAG GCACGTCAAGTCATTGTTTGTCTTGTGTTAAGTCACCCCAGGTGATTCTAAAGCCG AAGCTCTGCAAAGCACACCCTGAGAACACAGACAACCTCTGCTCTCGCTCTTGACAC TTCAGGTGCAAACACTTGTCTAACTGCTCTCAAACCTACGCATGTGAGAACATCAC [T, C] GTGAGAGCTTATTGAAACTGATTGCGGACCCATACCTAGAGGGCTGATTCTATAGGT CTGAGGTAAAGGCCAAGAATTGCTATTTGCTATTGCTTTCTTTCTTTCTTTCTT TTTTTTTTTTTGAGATGAAGTCTCACCTGTGCCAGACTGGAGTGAGTGGCATGA TCTCAGCTACTGCAGCCTCTGCCCTGGTTAAAGCGATTCTCCCCACACCCAGACC CGCTCCTGAGTAGCTGGGATTACAGGTGCCGCCACATGACTAGCTAACGTTGTATT [C, T] TTTTTTTTTTTGAGATGAAGTCTCACCTGTGCCAGACTGGAGTGAGTGGCAT GATCTCAGCTACTGCAGCCTCTGCCCTGGTTAAAGCGATTCTCCCCACACCCAGA CCCGCTCTGAGTAGCTGGGATTACAGGTGCCGCCACATGACTAGCTAACGTTGTATT TTTTAGTAGAGACGGGGTTTCAACATGTTGCCAGGCTGGCTCAAACCTGACCTCA GGTGTACCACTCACCTCAGCCTCCAAAGGTCTGGGATTACTGGGTGAGGCCACCGCGTG 18074
	TGCAGCCCTGCTGCCCTGGTTAAAGCGATTCTCCCCACACCCAGACCCGCTCTGAGT AGCTGGGATTACAGGTGCCGCCACATGACTAGCTAACGTTGTATTAGTAGAGAC GGGGGTTTCAACATGTTGCCAGGCTGGCTCAAACCTGACCTCAGGTGATCCACTCA

FIGURE 3, page 12 of 14

	CCTCAGCCTCCCAGGTCTGGGATTACTGGTGTGAGCCACCGCGTGCAGGCCAGAATTG CATTCTAACAGTCCCAGGTATGCTGATGCTGTGGTCCAGGGACACACTTGAGAAC [T, C, A]
	GCTTGTTACTCAGGCATATGTGGACAGTAGCGTCATCTCACCTGGGAGCTTCCCTGCAG CATCTCAGGCCTTGCCTACACCTACCAGATCAGAATCTGCACTTTAATCTCAATCCCCGC GTGATTCTCATGCACCTGGAAAGTTGAGAAATATGACCTAGAGGAGCCGAATGTGAAA CCACTGGAGGCAGAGATAGATGGAGAAATATCTCTTCTCACGGATACTAAAGATGCAA CAAAAAGGGCTGACTCTGGGTGTGCACCCAGGTGGGCTGATGACCGAAAAGAGGCCA
19328	TGTGTGTGAGGCCGGGAGTGCAGCGAGCCCCGAAATTCTCAGCCTAGTCCCCGCCA CATAGCTAACAGTGGAGGGAGGAGGTGAGAAGGAGTCAGTGCCTCACCTCCGGTG GAGTACCTGCTCTCTGTAGTTCTGCTCTGGGAGCTTGCCTGCTTCACCTCTCC CTCCATCCCCCTCTCTCACAGGAAAAATTCACCTTAATATTGAAAGTTCTCTCTCA GCAAGTCCTCTCAGGCACCCACAGGCAAAAGGAAACTAAGCAGAGTTAGGGCTTCCA [G, T] GCCTAGCCAACCTACACGACTCTCCTCTTGCTTCCCTAACGAAACCAGCGAAGGGCAGCGT GGGTTCCAGCATAGATGGACCTGTGTTGGAATCTCTGCACGTGCTGCTGACCCCTGGCT AGCCATTGACCTCTCTGAGCCCTGTTCTTCCACTAGGCTCTGAGGGCAGGGCC ATGTCCTTTTCACTGCTCTGCACTGAGCAGTGTGCAGGGCACATAGGAAGTTCCA TAAATGTTTGTGGGATAAAAGGAAATAAAACCTTCTCTCTGTCCCCCTGTGATGGC
19570	AAAGTCCTCTCAGGCACCCACAGGCAAAAGGAAACTAACGAGAGTTAGGGCTTCCAGG CCTAGCCAACCTACACGACTCTCCTCTTGCTTCCCTAACGAAACCAGCGAAGGGCAGCGT GGTTCCAGCATAGATGGACCTGTGTTGGAATCTCTGCACGTGCTGCTGACCCCTGGCTA GCCATTGACCTCTCTGAGCCCTGTTCTTCCACTAGGCTCTGAGGGCAGGGCC TGTCTTTTCACTGCTCTGCACTGAGCAGTGTGCAGGGCACATAGGAAGTTCCAT [A, G] ATATGTTGTGGGATAAAAGGAAATAAAACCTTCTCTCTGTCCCCCTGTGATGGCTT TGCACAAGGCACTGCTCTGGCCAGGTTTGCTAGGCTAGTGTGAGGATAAACCAAGGTATA TTACAAATTGGAGAAAATTCTCGTTCTTCTGGAAGAAGGTGCTGTATCATGAAACAAG AATGTCCTGATTCCCTCTATGCCAGGTACTGGGAGAAACAGGTGCTGATAACCGTTG ATCCAGGCAGAAATAAGCATACTCCTGCTTCCAAAGGCCATGCTTCTCTCTCC [C, T] AGAGGGAGGAGATAAGCTCAGGCCATGACACCAAGCCATGCTCTCTAGTCACCCACA GTGCTCTATGGTAATGAGTCAGTCAGCAACCAAGACGCTGAGGCCCTCTTTGGT GAGCCTGGCTGGGTGCTGAGGCTTCAGGTACAATCATGGGTTGGAAGAGGCCCTCTC TCTCCACAGTCTGGCACTATGACCCCTCTGGTTATTAAACAGGCAAAGAGAGAGAGGGA AGAAAAGCAGGCAAATAATGTTGGGTGCTATTCTAGAGATTAGAATTTCAGGAAGGATAA
20892	CCTTGGATGAAGAACGCTGGGACTCTTGCTTCTCTTCCCTCCGCAGTGACATGCCAT GCCATGCCACTGCCCTTCATCTGGCCTATGACAGTCAGTCAGTCATAAGCACCCGCATGTAC CCGGCCCTGCACTAGCTCATGACAGCTGAGTCAGTCATTGGCCAGGTGCTGTATCTCATCC GGCCTCTCAGCAACCCCTGAGATACTGGTAATGTCCTGATGAAGATATTACTGAGG CAGAAATGGACGCTCAGTGAAGCAAGGTGCTGATGTTAGCAATGAGCTATGAGTGGC [C, T] AGAGGGAGGAGATAAGCTCAGGCCATGACACCAAGCCATGCTCTCTAGTCACCCACA GTGCTCTATGGTAATGAGTCAGTCAGCAACCAAGACGCTGAGGCCCTCTTTGGT GAGCCTGGCTGGGTGCTGAGGCTTCAGGTACAATCATGGGTTGGAAGAGGCCCTCTC TCTCCACAGTCTGGCACTATGACCCCTCTGGTTATTAAACAGGCAAAGAGAGAGAGGGA AGAAAAGCAGGCAAATAATGTTGGGTGCTATTCTAGAGATTAGAATTTCAGGAAGGATAA
26465	TTCTCTGACCCCTCCCCCTCCGGTGCCTTCGTATCAAAGCTCTCAAACCCCGTCCCCCG TGTGTCTGCTGTGAGCTCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT CCCCAGGCTCGTGGGAGGTTCCACCCCTGGATCCAGGAAGAACCCCTCCACCCCTGCCCTCGT GGGTGGGCCAAAGGCTACAGGGTGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT TGCCATGGGCCTGCCCTCCCCAGTGACCTGCGAAAGTGGAGGCATCGAGGTAGGAGGGAAAC [G, A] GCAACCAGGGAGTCCTCGAGCCTGGGCTGCCCTACCTCTACCCATTCCCCGACCAGAGC TTTGCCTTGTGGCTGCCCTGCCCTCTTGGGAACTGAGCTCAGAGGCAGGTGCT TCAGAGAAGGAAACAAATGAGGGGGCAGGGATAAAAAGTCACCTCATTCTACCT CCCCATGCACTGAACACAATTCTCTCACCTGGCTCCAAATTAAAGATGTGGACCA AGGCCTGTGGGTACTCCAGGGGCAAGGAGAGGCCCTGGGTGAGTGCACACTGTCAGGCCAA
26472	ACCCCTCCCCCTCCGGTGCCTTCGTATCAAAGCTCTCAAACCCCGTCCCCGTGTGCTC TGCCTGCTGTGAGCTCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT TCGTGGGAGGTTCCACCCCTGGATCCAGGAAGAACCCCTCCACCCCTGCCCTCGTGGGTGG CCAAAGGCTACAGGGTGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT GCCCTGCCCTCCCCAGTGACCTGCGAAAGTGGAGGCATCGAGGTAGGAGGGAAACGGCAACC [A, G]

FIGURE 3, page 13 of 14

